0.5
6
5

THE UNIVERSITY
OF MICHIGAN
DEC 18 1961
MEDICAL
LIBRARY

# Journal of Speech and Hearing Research

VOLUME 2 December 1959 NUMBER 4

A JOURNAL OF THE
AMERICAN SPEECH AND HEARING
ASSOCIATION

## *General Information*

The American Speech and Hearing Association publishes the *Journal of Speech and Hearing Research*, the *Journal of Speech and Hearing Disorders*, *Monograph Supplements* and the *Employment Bulletin*.

The *Journal of Speech and Hearing Research* contains papers concerned with theoretical issues and research in the general area of speech and hearing disorders and related problems. It is published annually in March, June, September, and December and is printed by the Interstate Printers and Publishers, Inc., at 19-27 North Jackson Street, Danville, Illinois. Second-class postage is paid at Danville, Illinois.

*Subscription rates* are $5.00 a year in the United States; $5.50 a year outside the United States; $1.35 the single copy. Subscription orders and business correspondence should be sent to Kenneth O. Johnson, Business Manager, Journal of Speech and Hearing Research, American Speech and Hearing Association, 1001 Connecticut Avenue N.W., Washington 6, D. C.

*Manuscripts* and editorial correspondence should be sent to Dorothy Sherman, Editor, Journal of Speech and Hearing Research, Department of Speech Pathology and Audiology, East Hall, University of Iowa, Iowa City, Iowa. The Association reserves the right to keep any manuscripts submitted. It cannot assume responsibility for loss of manuscripts. No manuscript which has been printed elsewhere can be considered for publication in the *Journal of Speech and Hearing Research*. Before submitting manuscripts, authors are advised to consult *Information for Contributors to the Journal of Speech and Hearing Research* (*J. Speech Hearing Res.*, 1, 1958, 94-96). Copies of this note and its supplement, *Examples of Manuscript Form*, may be obtained by writing the editor.

*Permission to quote* a portion of any article in this issue of the *Journal of Speech and Hearing Research* is hereby granted if similar permission is obtained from the author (authors) and if credit is given the author (authors) and this Journal. An author (authors) or the sponsor of research may reproduce his (their) article in full, crediting this Journal; this permission is not assignable.

Journal
of Speech and Hearing
Research

December 1959 VOLUME 2 • NUMBER 4

Nonrandom Sources of Variation in Vowel Quality..... 305
WILLIAM R. TIFFANY

Articulation Skills, Velopharyngeal Closure, and
Oral Breath Pressure of Children with Cleft Palates..... 318
DUANE C. SPRIESTERSBACH AND GENE R. POWERS

Calling Attention to Stuttering..... 326
M. E. WINGATE

Post-exposure Residual Effects of Low-Level Noise..... 336
WILLIAM J. TRITTIPOE

Factors Related to Lipreading..... 340
AUDREY ANN SIMMONS

Intelligibility and Associated Physiological Factors
of Cleft Palate Speakers..... 353
JOANNE D. SUBTELNY AND J. DANIEL SUBTELNY

Effect of Loudness Recruitment on Delayed Speech Feedback..... 361
EARL R. HARFORD AND JAMES F. JERGER

Language Skills of Profoundly Deaf Adolescent Children..... 369
SIDNEY GODA

Language Skills of Male and Female Kindergarten Children..... 377
HARRIS WINITZ

Relationships Between Language and Nonlanguage
Measures of Kindergarten Children..... 387
HARRIS WINITZ

Index to Volume 2, 1959..... 392

# The American Speech and Hearing Association

**OFFICERS**

President
George A. Kopp, Ph.D.
Wayne State University

Executive Vice-President
Jack Matthews, Ph.D.
University of Pittsburgh

Vice-President
Miriam D. Pauls, Ph.D.
Johns Hopkins Medical Institute

Editor of the Association
Wendell Johnson, Ph.D.
University of Iowa

**OFFICERS-ELECT**

President-Elect
Stanley Ainsworth, Ph.D.
University of Georgia

Vice-President-Elect
Jack L. Bangs, Ph.D.
Houston Speech and Hearing Center

**COUNCIL**

The Officers and the following Councilors:

Jon Eisenson, Ph.D. (1959)
Ira J. Hirsh, Ph.D. (1958-60)
Ruth B. Irwin, Ph.D. (1956-59)
James F. Jerger, Ph.D. (1959-61)
Herold S. Lillywhite, Ph.D. (1958-59)
Hayes A. Newby, Ph.D. (1957-59)
Wilbert L. Pronovost, Ph.D. (1958-60)
Sylvia O. Richardson, M.D. (1957-59)
Dean E. Williams, Ph.D. (1959-61)

**EXECUTIVE SECRETARY**
Kenneth O. Johnson, Ph.D.

# The Journal of Speech and Hearing Research

**EDITOR**
Dorothy Sherman, Ph.D.

**ASSISTANT TO THE EDITOR**
Dorothy W. Moeller

**ASSISTANT EDITOR**
Frederic L. Darley, Ph.D.

**STATISTICAL CONSULTANT**
Leonard S. Feldt, Ph.D.

**ASSOCIATE EDITORS**
Oliver Bloodstein, Ph.D.
Arthur S. House, Ph.D.
James F. Jerger, Ph.D.
D. E. Morley, Ph.D.
Hildred Schuell, Ph.D.
William R. Tiffany, Ph.D.
John C. Webster, Ph.D.
Joseph M. Wepman, Ph.D.

**DEPARTMENT EDITORS**
Ernest H. Henrikson, Ph.D.
*Book Reviews*

Martin F. Palmer, Sc.D.
*Records*

**BUSINESS MANAGER**
Kenneth O. Johnson, Ph.D.

APPLICATIONS FOR MEMBERSHIP SHOULD BE ADDRESSED TO THE EXECUTIVE SECRETARY

# Nonrandom Sources of Variation in Vowel Quality

WILLIAM R. TIFFANY

It is well known that repeated utterances of the same vowel phoneme by a speaker or speakers will result in variations in observed phonetic quality as well as in measurable acoustic characteristics. Whereas much of this variation may be of an essentially random nature, certain nonrandom sources of variation have been discussed. It is well known, of course, that the phonetic context will influence the nature of the vowel. This was pointed out by Black (*1*) and others early in the history of acoustic phonetics. Even where phonetic context is held constant, however, and when the allophone is the unit studied, important inter- and intrasubject variations have been observed (*2, 3, 4*).

It is generally acknowledged that the production of vowel sounds by an individual is not a random process and that the values of the acoustic measurements of the vowels are therefore not distributed in random order. The data found by Peterson and Barney (*2*) were said to reveal that 'both the production and the identification of vowel sounds by an individual depend on his previous language experience.'

In their reference to the 'previous language experience' of the individual, Peterson and Barney were primarily concerned with the dialect influences, but did not specifically exclude other aspects of language experience which might not normally be thought of as dialectal in nature.

Among the possible sources of nondialectal, nonrandom variation in a single speaker's repeated utterances of an allophone, or among the utterances of many speakers, at least one has already been the subject of study. The influences of sex and age, as they govern resonator size, are well known. The data of Potter and Steinberg (*4*) and others make clear the relative nature of the formant patterns and appear to show conclusively that although formant ratios remain fairly stable the formant positions vary systematically with age and sex.

Except for normative studies which have contributed information concerning sex and age differences and studies which have investigated the language variable, there has been little attempt to identify the sources of nonrandom variations in vowel quality among and within subjects.

One possible source of variation in vowel quality which has not as yet been systematically studied is that contributed by speech training, or, what is presumed to be a result of this, speech excellence. Although it is difficult at first hand to see how this is a variable which is distinct from that of speech experience or dialect, an a priori case can be made for such a position.

To see how such a case might be made it is first necessary to admit a

William Tiffany (Ph.D., Iowa, 1951) is Associate Professor of Speech, University of Washington, Seattle.

factor of vowel intelligibility which is dependent upon the degree to which the several vowels of a given vowel system are mutually distinctive. High intelligibility might then be considered some function of the number of difference limens separating the vowels, which would in turn depend upon the separateness of the main resonance regions of the contrasted vowels. In terms of the first formant and the second formant as plotted on the familiar acoustic vowel diagram, the posited vowel intelligibility would thus be a function of the size of the vowel diagram, as opposed to its shape. In this view the larger the vowel diagram the more intelligible would be the speech. It seems possible that this variable of vowel diagram size would be in some degree independent of differences conferred by a dialectal factor. It would seem not unreasonable to suppose that increased vowel intelligibility, as defined by vowel diagram size, would be associated with speech excellence or with speech training.

Another possible source of nonrandom variation in vowel utterance, which has as yet received no systematic study, is the kind of stress which commonly results from forcefulness of utterance. Previous studies have shown that there is reason to believe that the several factors which enter into stress, such as length, fundamental pitch change, and intensity, are normally independent of vowel formant position, but it would seem possible nevertheless that such stress might be accompanied, in certain lawful ways, by concomitant changes in resonance. That is, it is conceivable that a part of the stress of a vowel might result from a typical alteration in its quality. Can a vowel be stressed by making it more intelligible? For example, is a stressed [i] more '[i]-ish' than an unstressed [i]? There is, of course, much good experience and evidence to show that this indeed is the case, and there is recognition of the obscuring of the vowel which occurs when a syllable loses its accent. Kenyon (*5*, p. 90) points this out when he states that 'It is characteristic of English . . . that the vowels of unaccented syllables have gradually obscured to a sound quite different in resonance. . . .' The degree and the direction in which this obscuring operates, the nature of such obscuring, and the presence of obscuring in normal variations in sense stress (as opposed to syllabic stress) have not been investigated systematically as a function of formant frequency location.

It was the purpose of this investigation to assess the influence on vowel resonances of two factors: (a) the factor of speech excellence, as this might be assumed to be a result of speech training and experience, and (b) the factor of sense stress. In addition, a third factor, that of the presence or absence of a consonant context, also was investigated, for what light it might throw not only on differences between trained and untrained speakers, but also on the general question of the so-called isolated phoneme.

Put in more concrete terms, this study makes an attempt to determine what, if any, significant acoustic differences might be found (a) between the vowel utterances of trained speakers and a matched group of untrained speakers of the same dialect characteristics, (b) between stressed and unstressed vowels in the same context,

and (c) between vowels in context and vowels in isolation.

Not only was it hoped that data could be provided pertinent to the problems of articulation improvement, but also it was hoped that light could be thrown on other problems concerned with the stability of the vowel resonance patterns. For example, do the patterns, as represented by the acoustic vowel diagram, vary in certain predictable ways among and within speakers of the same language habits? Can vowel intelligibility be thought of as related to the general size of the vowel diagram, as opposed to its shape? Are stressed vowels merely louder, longer, and higher? Or do they characteristically occupy a different position on the vowel diagram? Are better speakers better because of prosodic factors, semantic factors, and consonant articulation factors only, or do they perhaps also use vowel sounds which are less variable or of a different resonance? What relation does the quality of the isolated vowel phonation have to the quality of the vowel in context? These are some of the more specific questions which this research seeks to answer.

## *Experimental Method*

The design of this experiment involved the measurement of the acoustic characteristics of each of 10 vowels [i], [ɪ], [e], [ɛ], [æ], [o], [ʊ], [u], [ɑ], and [ʌ] and two diphthongs [aɪ] and [aʊ] under three different conditions of speaking for two different groups of adult General American speakers.

*Speaking Conditions.* There were two different conditions of sense stress for the 12 different vowel and consonant sounds in a [h__d] context. Contrastive sense stressing was obtained by devising six more or less sensible sentences, each of which contained two different [h__d] words. These sentences were presented to the reader in pairs, with different words underlined in each member of the pair, in such a way that if read properly the two readings would produce both a stressed and an unstressed reading of the same word in the same context. One such pair was:

They hid under the hood of the car.

They hid under the hood of the car.

The instructions to the readers were, 'Read the following in a natural tone, but making it very clear that the meaning differs in each pair according to the word stressed (indicated by the underlining).' The six sentences employed in the readings were:

a. But he'd never had the warning.

b. I don't know who'd have thought to hide there.

c. They hid under the hood of the car.

d. Now how'd you like to be the head boss.

e. The farmer hayed the field and hoed the potatoes.

f. The old hod belonged to Mr. Hud.

This technique produced for analysis 12 vowels with phonetic and sentence contexts constant, but with two different conditions of sense stress. No instructions were given as to how stressing was to be carried out. Nearly all

speakers, trained and untrained, were able to carry out the instructions without difficulty. In the case of obvious errors the sentence was re-recorded on the spot under the same conditions and instructions.

A third condition of speaking was that of the isolated vowel phonation. In this condition, which followed the reading of the sentences, subjects were presented with a list of the 12 [h__d] words out of context and were told, 'For each of the following words read only the vowel sound.' The trained speakers had no difficulty doing this. Most of the untrained speakers also found it simple to follow these directions. Two of the speakers had to be trained, however. This training consisted simply of saying, 'No, what I mean is say the vowel all by itself. For example, for the word [hɝd] you would say only [ɝ].' After a few trials this was sufficient in all instances. No examples were given, of course, of the sounds under study.

*Speakers.* The trained-speaker group, 10 males, of this study was chosen from among the faculty and advanced graduate students of the Department of Speech, University of Washington. Although these people were not all of high artistic attainment or of beautiful voice quality, they demonstrated good, clear, articulate speech. They were all of the General American dialect and appeared to the experimenter to demonstrate no dialect peculiarities in their speech—at least insofar as the vowel enunciation was concerned.

The untrained speakers, 10 males, were students in a beginning speech course. They had as yet received no significant amount of formal phonetic training, and were all judged by the experimenter to be 'indifferent' to 'poor' with regard to what might be termed general articulation accuracy. All spoke with a General American dialect, however, and in this respect were not judged (by the experimenter) to be in any significant degree different from the trained speakers. In physical size they were approximately matched to the trained speakers, although there was no formal control placed upon this factor. The two groups of speakers differed most obviously from one another in amount of experience and training.

*Analysis of Recordings.* All phonations were recorded on an Ampex-600 tape recorder in a relatively dead quiet room. These recordings were then dubbed onto the recording tape of the Kay Electric Company Sonagraph for analysis.

An attempt was made to equate all vowels dubbed onto the Sonagraph recorder in terms of overall intensity as this may be determined from a VU meter. The vowels were dubbed at a level which would result in a −5 VU reading when the slow speed playback was monitored, regardless of whether the vowel was stressed, unstressed, or isolated. Dubbing was thus accomplished with considerable monitoring of the recording volume inasmuch as stressed vowels were, of course, considerably more intense than the unstressed.

*Acoustic Measures.* For each vowel spectrogram four acoustic measures were obtained, in the following way:

(a) A duration measure was made directly from the spectrogram. In general the duration of the vowel was defined as the duration of at least two

clear-cut formants in the appropriate frequency regions.

(b) A measure of the fundamental frequency was made by averaging the number of vertical striations of the broad-band spectrogram over a .05-sec interval in the mid-portion of the vowel. This, of course, substituted a single frequency measurement, presumed to be representative of that vowel because taken in its mid-portion, for the complex pattern of frequency change.

(c) and (d) Measures of the frequencies of formants one and two were made directly from the broad-band spectrogram at a point one-fourth the duration of the whole vowel. This point was chosen somewhat arbitrarily as representative of the whole vowel because perhaps least affected by onset or terminating influences.

For each of the two distinctive diphthongs a somewhat different method of measurement was used for formants one and two. Two sets of measures were made: one of the formant positions at the very onset of the diphthong and a second at its termination.

## *Results*

Tables 1 and 2 present the basic data of this experiment in terms of mean measurements for 10 subjects in each of two groups. Table 1 gives mean data for the duration and fundamental frequency measures. Table 2 gives both means and standard errors for the first formant and second formant measures, for trained and untrained speakers, and for the three conditions of phonation (stressed, unstressed, isolated).

Tables 3, 4, and 5 present a summary of the only statistical analyses thus far attempted. Although it may appear from Table 2 that all formant data could be subjected to an analysis of variance, several reasons, not excluding the inordinate amount of time involved, argue forcefully against this procedure. Tables 3, 4, and 5 present, instead, the difference scores obtained between certain pairs of formant means, with differences significant at or beyond the 5% level, as determined by a *t* test, indicated. Table 3 presents the difference scores comparing the trained with the untrained speakers. Table 4 compares stressed with unstressed phonations, and Table 5 compares the stressed with the isolated phonations.

It can be hypothesized that in the process of articulation it is the relation-

TABLE 1. Duration (in seconds) and frequency (in cycles per second) means for each of 10 vowels phonated by 10 trained and 10 untrained speakers. Frequency measures represent an average frequency over .05 sec taken from the center of the vowel phonation.

| *Vowel* | *Stressed* | | *Unstressed* | |
|---|---|---|---|---|
| | *Trained* | *Untrained* | *Trained* | *Untrained* |
| *(Duration)* | | | | |
| [i] | .163 | .159 | .096 | .123 |
| [ɪ] | .168 | .154 | .128 | .124 |
| [e] | .182 | .180 | .143 | .132 |
| [ɛ] | .107 | .124 | .100 | .107 |
| [æ] | .152 | .159 | .123 | .129 |
| [o] | .184 | .189 | .131 | .156 |
| [ʊ] | .147 | .135 | .092 | .101 |
| [u] | .171 | .178 | .119 | .111 |
| [ɑ] | .281 | .247 | .266 | .220 |
| [ʌ] | .188 | .170 | .177 | .156 |
| Mean | .174 | .170 | .138 | .136 |
| *(Frequency)* | | | | |
| [i] | 193 | 187 | 126 | 137 |
| [ɪ] | 162 | 175 | 121 | 132 |
| [e] | 183 | 175 | 132 | 146 |
| [ɛ] | 171 | 169 | 120 | 121 |
| [æ] | 204 | 186 | 135 | 134 |
| [o] | 152 | 154 | 134 | 138 |
| [ʊ] | 177 | 176 | 106 | 127 |
| [u] | 184 | 206 | 113 | 134 |
| [ɑ] | 121 | 153 | 83 | 111 |
| [ʌ] | 145 | 138 | 82 | 123 |
| Mean | 169 | 172 | 115 | 130 |

TABLE 2. Means and standard errors for the first two formant positions for 10 trained and 10 untrained speakers.

| *Vowel* | *Untrained* | | | | | | *Trained* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Unstressed* | | *Stressed* | | *Isolated* | | *Unstressed* | | *Stressed* | | *Isolated* | |
| | *M* | *SE* | *M* | *SE* | *M* | *SE* | *M* | *SE* | *M* | *SE* | *M* | *SE* |
| *(Formant 1)* | | | | | | | | | | | | |
| [i] | 395 | 20 | 400 | 15 | 350 | 10 | 375 | 13 | 395 | 9 | 300 | 7 |
| [ɪ] | 480 | 13 | 505 | 16 | 485 | 15 | 440 | 16 | 515 | 13 | 450 | 15 |
| [e] | 520 | 20 | 510 | 16 | 560 | 14 | 505 | 14 | 500 | 13 | 495 | 16 |
| [ɛ] | 595 | 17 | 615 | 15 | 600 | 21 | 590 | 14 | 620 | 13 | 615 | 20 |
| [æ] | 710 | 18 | 760 | 21 | 745 | 22 | 680 | 24 | 725 | 26 | 720 | 23 |
| [o] | 545 | 14 | 530 | 20 | 555 | 23 | 540 | 10 | 455 | 20 | 495 | 26 |
| [ʊ] | 560 | 16 | 570 | 11 | 540 | 16 | 505 | 9 | 530 | 15 | 480 | 17 |
| [u] | 450 | 18 | 485 | 15 | 435 | 21 | 410 | 14 | 440 | 24 | 360 | 12 |
| [ɑ] | 700 | 10 | 725 | 20 | 705 | 16 | 750 | 17 | 750 | 18 | 700 | 11 |
| [ʌ] | 680 | 17 | 740 | 19 | 680 | 21 | 675 | 15 | 730 | 19 | 680 | 8 |
| [aɪ] | 730 | 20 | 785 | 36 | 805 | 28 | 700 | 31 | 810 | 46 | 775 | 21 |
| [a͡ɪ] | 435 | 27 | 415 | 26 | 475 | 42 | 415 | 24 | 385 | 25 | 405 | 31 |
| [a͡ʊ] | 785 | 26 | 805 | 25 | 785 | 18 | 720 | 31 | 733 | 26 | 850 | 15 |
| [a͡ʊ] | 690 | 44 | 660 | 26 | 470 | 48 | 700 | 27 | 656 | 30 | 340 | 37 |
| *(Formant 2)* | | | | | | | | | | | | |
| [i] | 2170 | 66 | 2345 | 41 | 2305 | 36 | 2255 | 35 | 2355 | 27 | 2450 | 41 |
| [ɪ] | 2075 | 47 | 2215 | 22 | 2095 | 56 | 2155 | 42 | 2180 | 24 | 2150 | 36 |
| [e] | 2150 | 22 | 2175 | 25 | 2105 | 59 | 2190 | 50 | 2250 | 31 | 2185 | 58 |
| [ɛ] | 1930 | 52 | 1960 | 47 | 1970 | 54 | 1900 | 50 | 1985 | 42 | 2025 | 56 |
| [æ] | 1770 | 19 | 1845 | 39 | 1830 | 42 | 1785 | 27 | 1865 | 43 | 1955 | 55 |
| [o] | 935 | 36 | 940 | 54 | 935 | 55 | 900 | 15 | 805 | 32 | 765 | 18 |
| [ʊ] | 1300 | 48 | 1275 | 40 | 1210 | 46 | 1210 | 34 | 1140 | 33 | 1110 | 52 |
| [u] | 935 | 55 | 1025 | 34 | 1056 | 49 | 922 | 40 | 930 | 30 | 765 | 29 |
| [ɑ] | 1125 | 40 | 1150 | 33 | 1145 | 35 | 1205 | 17 | 1235 | 25 | 1110 | 47 |
| [ʌ] | 1380 | 60 | 1370 | 37 | 1315 | 40 | 1385 | 22 | 1370 | 15 | 1265 | 33 |
| [aɪ] | 1335 | 36 | 1230 | 28 | 1295 | 31 | 1340 | 46 | 1275 | 46 | 1265 | 25 |
| [a͡ɪ] | 1920 | 40 | 2045 | 50 | 2135 | 74 | 1995 | 48 | 2100 | 53 | 2310 | 48 |
| [a͡ʊ] | 1320 | 55 | 1300 | 30 | 1345 | 60 | 1295 | 32 | 1300 | 41 | 1350 | 32 |
| [a͡ʊ] | 1210 | 63 | 1090 | 54 | 695 | 67 | 1220 | 50 | 1144 | 38 | 525 | 51 |

ship among the resonances of the several vowels which determines recognition, rather than the resonance frequency per se. Figures 1 and 2 are presented to show these relationships graphically on a conventional acoustic vowel diagram with the first formant plotted against the second formant and with coordinates arranged according to the mel scale. In Figure 1 the formant coordinates for the three conditions are compared for trained speakers only. Figure 2 gives the same data for the untrained speakers.

*Frequency and Duration.* The frequency and duration measures of Table 1 indicate clearly that the stressed phonations were indeed given more prominence by the speakers, at least insofar as the measures employed here are concerned. Duration averages for stressed vowels were about .17 sec and for unstressed less than .14 sec. The differences are clearly statistically significant. It is also clear that stressed vowels were significantly higher in pitch than the unstressed, with stressed vowels averaging about 170 cps and unstressed

TABLE 3. Mean difference scores* showing the comparison between trained and untrained speakers on the basis of formant frequency measures, all measures in cycles per second.

| *Vowel* | *Stressed* | | *Unstressed* | |
|---|---|---|---|---|
| | *Formant 1* | *Formant 2* | *Formant 1* | *Formant 2* |
| [i] | 5 | −10 | 20 | −85 |
| [ɪ] | −10 | 35 | 40 | −80 |
| [e] | 10 | −75 | 15 | −40 |
| [ɛ] | − 5 | −25 | 5 | 30 |
| [æ] | 35 | −20 | 30 | −15 |
| [o] | 75† | 135† | 5 | 35 |
| [ʊ] | 40† | 135† | 55† | 90 |
| [u] | 45 | 95† | 40 | 13 |
| [ɑ] | −25 | −85† | −50† | −80 |
| [ʌ] | 10 | 0 | 5 | − 5 |
| [a͡ɪ] | −25 | −45 | 30 | − 5 |
| [ā͡ɪ] | 30 | −55 | 20 | −75 |
| [a͡ʊ] | 72 | 0 | 65 | 25 |
| [ā͡ʊ] | 4 | −54 | −10 | −10 |

*Difference scores are obtained by substracting the mean for the 10 trained speakers from the mean for the 10 untrained.
†Statistically significant at or beyond the 5% level.

about 122 cps. Such findings are, of course, in agreement with many previous studies on the acoustical nature of accent.

On the whole, pitch and duration trends were similar for both trained and untrained subjects. One interesting pitch difference, however, is clearly evident from Table 1. It will be noted that whereas both groups used about

TABLE 4. Mean difference scores* showing the comparison between stressed and unstressed vowels on the basis of formant frequency measures, all measures in cycles per second.

| *Vowel* | *Untrained Speakers* | | *Trained Speakers* | |
|---|---|---|---|---|
| | *Formant 1* | *Formant 2* | *Formant 1* | *Formant 2* |
| [i] | 5 | 175† | 20 | 100† |
| [ɪ] | 25 | 140† | 75† | 25 |
| [e] | −10 | 25 | − 5 | 60 |
| [ɛ] | 20 | 30 | 30 | 85† |
| [æ] | 50† | 75 | 45† | 80† |
| [o] | −15 | 5 | −85† | −95† |
| [ʊ] | 10 | −15 | 25 | −70 |
| [u] | 35† | 90 | 30 | 8 |
| [ɑ] | 25 | 25 | 0 | 30 |
| [ʌ] | 40† | −10 | 55† | −15 |
| [a͡ɪ] | 55 | −105† | 110† | −65† |
| [ā͡ɪ] | −20 | 125† | −30 | 105† |
| [a͡ʊ] | 20 | −20 | 13 | 5 |
| [ā͡ʊ] | −30 | −120† | −44 | −76 |

*Difference scores are obtained by subtracting the mean for the unstressed vowel from the mean for the stressed.
†Statistically significant at or beyond the 5% level.

TABLE 5. Mean difference scores* showing the comparison between stressed and isolated vowels on the basis of formant frequency measures, all measures in cycles per second.

| *Vowel* | *Untrained Speakers* | | *Trained Speakers* | |
|---|---|---|---|---|
| | *Formant 1* | *Formant 2* | *Formant 1* | *Formant 2* |
| [i] | 50† | 40 | 95† | –95† |
| [ɪ] | 20 | 120 | 65† | 30 |
| [e] | –50† | 70 | 5 | 65 |
| [ɛ] | 15 | –10 | 5 | –40 |
| [æ] | 15 | 15 | 5 | –90 |
| [o] | –25 | 5 | –40 | 40 |
| [ʊ] | 30 | 65 | 50 | 30 |
| [u] | 50 | –31 | 80† | 165† |
| [ɑ] | 20 | 5 | 50 | 125† |
| [ʌ] | 60† | 55 | 50† | 105† |
| [a͡ɪ] | –20 | –65† | 35 | 10 |
| [a͡ɪ] | –60 | –90 | –20 | –210† |
| [a͡ʊ] | 20 | –45 | –117† | –50 |
| [a͡ʊ] | 190† | 395† | 316† | 619† |

*Difference scores are obtained by subtracting the mean for the isolated vowel from the mean for the stressed.
†Statistically significant at or beyond the 5% level.

the same pitch for the stressed vowels the trained speakers were consistently lower in pitch on the unstressed. This might suggest that appropriate unstressing is more a mark of the successful speaker than is appropriate stressing.

*Formant Location as a Function of Speech Training.* As Tables 2 and 3 show, there are some indications that speech training is a factor influencing vowel quality, although there are obvious and complex interactions between this factor and the vowels and stresses.

The data in Table 2 seem to indicate some slight tendency for consistency and training to go together, at least as that consistency is measured by an among-speakers variability. If only the context vowels are taken into account, then out of a total of 56 standard error measurements (14 vowels and diphthongs, 2 stresses, 2 formants), the untrained speakers show the greater variation in 33 cases, the trained in 23. This is a result which presumably could be due to chance alone about once out of four times. The trend is clearest in the case of the second formants for the 'pure' vowels, where 15 out of 20 measures are less variable for the trained speakers than for the untrained.

As indicated in Table 3, mean differences between trained and untrained speakers show slight tendencies for vowel quality to vary in consistent and lawful fashion. Differences between groups are significant for four of the stressed vowels and two of the unstressed. In general, first formant positions were lower for the trained speakers. The most obvious exception is in the case of [ɑ]. Of 28 first formant differences there are only six exceptions to the rule of a lower first formant for trained speakers than for untrained. Second formant measures give a different picture. There is a tendency for the back vowels (except for [ɑ]) of trained speakers to have lower second formants, while front vowels are, for

the most part, higher in second formant postition (8 out of 10 measures). All of the statistically significant differences are for the back vowels.

The possible significance of the above differences is more evident on comparison of Figures 1 and 2, in which the interaction between stressing, training, and formant ratios can be better evaluated. The comparisons are highly complex and at this point the reader is left to formulate his own hypotheses concerning the differences between Figures 1 and 2. These figures will be discussed in the following paragraphs.

*Formant Location as a Function of Condition of Phonation (Stressed, Unstressed, Isolated).* There does not appear to be any well defined tendency for variability of formant position and stress to go together. A slight tendency does exist for the second formant measures to be more variable in unstressed vowels. However, this tendency is somewhat weak and in only a very few cases do the differences appear to approach statistical significance.

The relationship between formant frequency means and stress is found to be very complex and to involve important interactions between stress and training. There are, however, certain interesting trends to be observed in the data of Tables 2 and 4. Of the 40 formant differences obtained between stressed and unstressed measures, 31 show a higher formant position for the stressed phonation. Of all of the significant differences found—of which there were 13 for the vowels—only two (both for [o]) showed a reversal of this trend. Of all of the vowels measured only [e], [ʊ], and [ɑ] failed to show some significant differences, for either the trained or the untrained groups, for either the first formant or the second formant. There appears to be a tendency for stressing differences to be greater for the trained than for the untrained speakers.

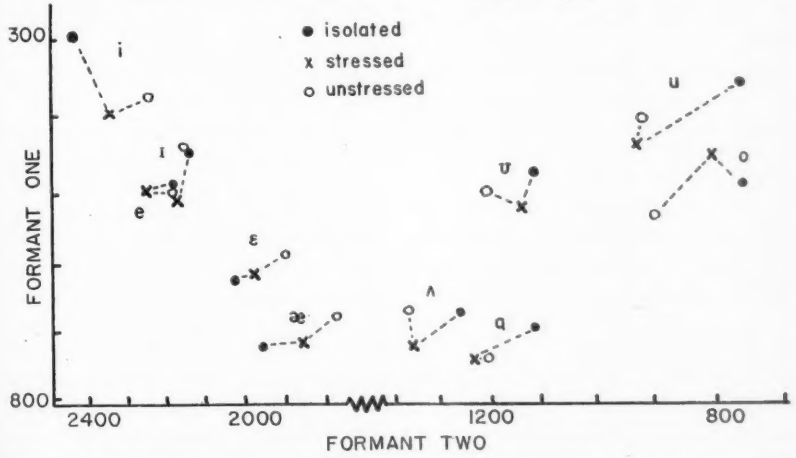



FIGURE 1. Acoustic vowel diagram for isolated, stressed, and unstressed vowels as produced by the trained speakers. Coordinates are based upon the average of 10 speakers. The scales are mel scales, with a mid-section compressed for the formant two scale.

Changes of considerable magnitude are also apparent for the diphthongs. Second formant changes were large and significant for both the beginning and ending of [aɪ] for both trained and untrained speakers and for the ending of [ɑʊ] for untrained speakers.

Table 5 shows the magnitude and the significance of differences between the stressed and isolated phonations. Tendencies are seen to vary markedly but to be generally in the direction of lower formants for the isolated vowels (32 out of 40 comparisons). There are slight tendencies to find greater differences for trained speakers than for untrained speakers in this respect. Differences are particularly distinctive for the high front vowels, the low back vowels, and the diphthongs. Attention should be drawn especially to the very large differences between the context and the isolated [aʊ]—pronounced as a rather extreme diphthong in isolation but almost as a vowel [a:] in speech.

Figures 1 and 2 present the complex data of the tables in a somewhat more comprehensible form, on a vowel diagram with the first formant plotted against the second formant. They are offered here without any attempt at statistical analysis, but as a means whereby the reader may form his own hypotheses regarding the influence of stress and training on vowel quality. There does appear in these figures a tendency for the vowel diagram to grow smaller from isolated vowel to stressed vowel to unstressed vowel. Vowels seem to move toward a neutral, or at least a central, point on the vowel diagram as they lose energy in context. This is a somewhat weak trend, of course, but it may be seen operating to some extent for the [i], [ɛ], [æ], [u] and [o]. The trend is somewhat less obvious for [ʌ] and [ɑ], as one might expect, and is to some extent reversed for [ɪ] and [e]. Incomplete as the data are it does appear possible

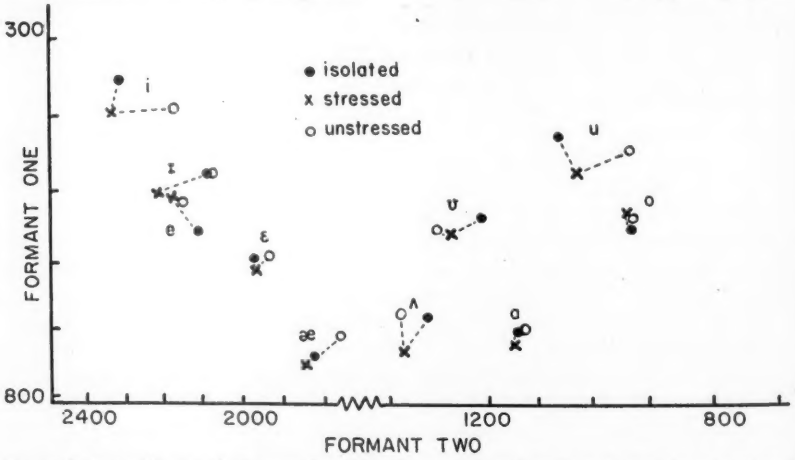



FIGURE 2. Acoustic vowel diagram for isolated, stressed, and unstressed vowels as produced by the untrained speakers. Coordinates are based upon the average of 10 speakers. The scales are mel scales, with a mid-section compressed for the formant two scale.

to hypothesize that for trained speakers, at least, stressing and context influence vowel diagram size, and probably more than they influence vowel diagram shape.

Figure 2, the diagram for the untrained speakers, by no means shows the same extent of diagram shift as does Figure 1. To be sure, when only stressed and unstressed vowels (in context) are examined the same tendency is observed, though with less strength ([ɑ], [o] and [u] are apparent reversals). The isolated phonations appear not to have exactly the same relationship to the context phonations as they did for the trained speakers. For the trained speakers the isolated phonation seems more often to appear as a kind of exaggeration of the normal vowel quality, whereas for the untrained speakers it appears more as merely a kind of representative of the average vowel quality.

## *Discussion*

It was the stated purpose of this study to assess the influence on vowel resonance of speech excellence, sense stress, and vowel isolation. Significant differences were found, for some vowels at least, indicating that these factors may be of importance in determining the precise vowel color for any given context. Although differences were complex and contradictory in some ways it does seem possible to hypothesize a kind of vowel intelligibility which may be related to the size, as opposed to the position or shape, of the vowel diagram. With regard to the questions asked in the introduction, stressed vowels may not be merely longer, higher, and louder than unstressed. They may also be different in terms of vowel resonance patterns as well. The differences, moreover, appear to be differences of pattern, involving the vowel diagram as a whole, such that a stressed vowel, for example an [i], is not only longer, louder, and higher, but is also more '[i]-ish' than an unstressed vowel.

It may be possible now to fit the present data into a three-part hypothesis which would account for most of the nonrandom sources of variation in vowel recognition. In terms of a bidimensional model, the vowel system can be conceptualized as a series of first formant (F1), second formant (F2) coordinates. Vowel recognition, in this system, may be thought of as a process of discriminating among these points. This job of discrimination will be aided by the fact that for any one person repeating a single allophone under constant conditions, variations of these points will be essentially random in nature and within rather well defined limits.

Speech recognition depends, however, upon ability to understand more than a single person and condition and therefore upon ability to adapt to other than random sources of variation. In two dimensional terms this and previous research indicate that recognition may depend upon ability to adapt to at least three kinds of nonrandom variation, as illustrated in Figure 3.

First illustrated, in Figure 3A, is the situation in which the only sources of variation would be random. The extent of such variation would determine the relative reliability of the coordinates which fix the corners of diagram A, as indicated by the large dotted

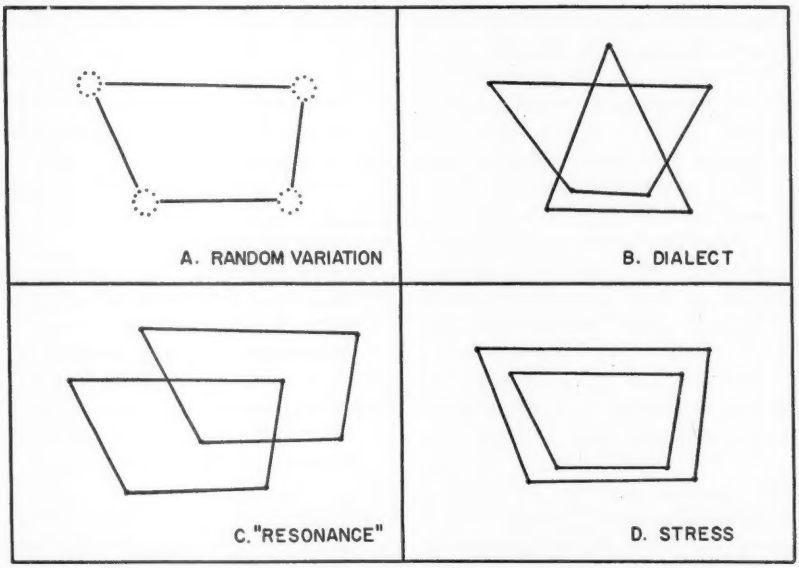



FIGURE 3. Graphic illustration of four kinds of vowel variation given in terms of the acoustic formant one, formant two vowel diagram. Ordinate is formant one; abscissa is formant two.

circles, but it would not determine the relationship among the circles.

The second kind of vowel variation, illustrated in 3B, is that of dialect, perhaps the most difficult kind to which the listener must adapt. The figure illustrates the vowel coordinates and relationships for two different (hypothetical) languages or dialects.

The third kind of vowel variation is suggested in 3C which illustrates the variation encountered between adults and children, or as a result of changes in overall resonator size. In this case the vowel relationships are relatively constant but the positions of the vowel diagrams within the F1-F2 space are shifted. Success in understanding the speech of both men and children testifies to ability to make such an adaptation; that is, as long as vowel diagram shape remains constant, its position may change, within limits, without interfering with recognition. The limits of such change have not yet been well defined.

The fourth kind of variation is that which has to some extent been demonstrated in the present study. This is illustrated in Figure 3D. Here the vowel diagram shape remains the same and the location in the F1-F2 space is roughly equivalent, but the size has changed. The tendency for the stressed vowels to produce a diagram which fills more of the available F1-F2 space, together with the tendency for trained speakers to show somewhat more extreme vowel coordinate positions when producing isolated vowels, would lend support to 3D as illustrative of an important source of variation.

One possible hypothesis arising from a consideration of this fourth kind of variation is the following: other things being equal, the larger the vowel diagram size, the better the vowel intelligibility. How might such an acoustic hypothesis be applied to speech training problems and what would it mean in terms of physiological phonetics? Perhaps it may be said that for increased intelligibility the tongue should be higher and more forward for high front vowels, and lower and more to the back for the back vowels. Perhaps it should simply be said that more extreme articulatory positions are characteristic of more careful articulation or of stress. The present investigation has, of course, not yet proved the point.

This study has, it would seem, made it possible to advance some hypotheses in the area of vowel intelligibility. It is left for further study, however, to determine the limits within which vowel diagram size may be manipulated with good retention of intelligibility. As a matter of fact the whole question of the limits within which the F1-F2 space may be manipulated with respect to size, position, and shape deserves further attention.

## *Summary*

Nonrandom sources of variation in vowel quality were studied by means of acoustic measures of duration, fundamental frequency, and frequencies of the first two formants of isolated vowels and of stressed and unstressed vowels in an [h—d] context for 10 trained and 10 untrained speakers. Although results indicated complex and variable relationships, there was some evidence that, in general, the differences among the several vowel resonances tended to be greater for the stressed than for the unstressed vowels, and greater for the trained than for the untrained speakers. There was also evidence that isolated vowels produced by trained speakers have more extreme articulatory positions than the same vowels in context.

A hypothesis was suggested to the effect that vowel intelligibility may be considered as a kind of function of vowel diagram size as well as of shape, location, and reliability.

## *References*

1. Black, J. W., The effect of the consonant on the vowel. *J. acoust. Soc. Amer.*, 10, 1939, 203-205.
2. Peterson, G. E., and Barney, H. L., Control methods used in a study of the vowels. *J. acoust. Soc. Amer.*, 24, 1952, 175-184.
3. Potter, R. K., and Peterson, G. E., The representation of vowels and their movements. *J. acoust. Soc. Amer.*, 20, 1948, 528-535.
4. Potter, R. K., and Steinberg, J. C., Toward the specification of speech. *J. acoust. Soc. Amer.*, 22, 1950, 807-820.
5. Kenyon, J. S., *American Pronunciation.* Ann Arbor: G. Wahr, 1946.

# Articulation Skills, Velopharyngeal Closure, and Oral Breath Pressure of Children with Cleft Palates

DUANE C. SPRIESTERSBACH

GENE R. POWERS

This study deals with children with cleft lips and palates and the relationships among their articulation skill, oral breath pressure, and velopharyngeal closure as revealed by lateral head x rays.

Oral pressure has long been recognized as one of the requisites in the production of speech sounds. Stetson and Hudgins (*12*), Hudgins and Stetson (*6*), Stetson (*11*), and Black (*1*) have studied the amount of breath pressure in the oral cavity of normal speakers during speech. They have shown that oral breath pressure is greater during the production of fricatives and plosives than it is either for other types of consonants or for vowels. They also have found that the pressure is greater during the production of voiceless consonants than during the production of their voiced cognates.

The pattern of articulation errors of individuals with cleft palates is of considerable interest in view of the findings summarized above. Investigators such as Bzoch (*2*), Counihan (*3*), McWilliams (*8*), Spriestersbach and others (*9*), and Starr (*10*) have found that individuals with cleft palates misarticulate with greatest frequency those sounds for which oral breath pressure is highest for normal speakers. These findings suggest that cleft palate speakers typically have difficulty building up enough air pressure in the oral cavity for the adequate production of speech. The most reasonable explanation for this lack of pressure is that there is an abnormal air leak through the velopharyngeal valve during articulation.

Films of lateral head x rays have been used extensively to assess the adequacy of velopharyngeal closure. Although x-ray films provide data of major diagnostic value, they deal with two spatial dimensions only and do not permit assessment of the volumetric dimensions of the oral and nasal cavities. The clinician needs to be able to augment the x-ray data with more direct evaluations

Duane C. Spriestersbach (Ph.D., University of Iowa, 1948) is Professor of Speech Pathology and Audiology, Department of Speech Pathology and Audiology and Department of Otolaryngology and Maxillofacial Surgery, University of Iowa. Gene R. Powers (M. A., University of Iowa, 1958) is a Research Associate, University of Iowa. This paper was given in part at the 1958 Convention of the American Association for Cleft Palate Rehabilitation, San Francisco. The investigation was supported in part by research grant M-1158 from the National Institute of Mental Health, Public Health Service.

of the functional adequacy of closure. Measurement of oral breath pressure is one way to provide this functional assessment.

### *Procedure*

The 103 subjects in the present study consisted of 38 girls and 65 boys ranging from five through 15 years of age. The mean IQ for the entire group as measured by the Wechsler Intelligence Scale for Children was 93.6. Eight of the group had socially-handicapping hearing losses, that is, an average loss of 25 db or greater in the better ear for the speech frequencies. Eight of the group had clefts of the lip only, 23 had clefts of the palate only, and 72 had clefts of both lip and palate. Of the 95 subjects with palatal clefts, seven had unrepaired cleft palates, 19 were wearing obturators, and 69 had had surgical closure of the palate. The physical management of the subjects had been carried out at various centers in four midwestern states.

Articulation skill was measured by the Templin diagnostic picture articulation test. This test elicits 176 word responses and allows for the evaluation of all of the speech sounds in a variety of phonetic contexts. Furthermore, normative data are available which make it possible to compare an individual's score on the test with the scores of physically normal children of the same age and sex. For the purposes of this study a ratio was computed between the individual's articulation score and the norm for his age and sex. The ratio was used in an attempt to minimize the effects of maturation by taking into account the normal pattern of development of articulation skills. A ratio of 1.00 was assigned if the performance of a given subject equalled the norm. Thus, the lower the ratio the less favorable was the individual's performance in comparison to the norm. This ratio will be referred to as the Templin ratio. The number of correct responses for 72 pressure sounds (36 plosives and 36 fricatives) elicited during the Templin test was used as a second articulation measure. The specific sounds were arbitrarily chosen but were uniform for all subjects. They were selected to sample all of the fricative and plosive sounds in the initial, medial, and final positions. Singles and blends were sampled equally for both fricatives and plosives.

Two measurements of oral pressure also were obtained. One was a measure of vital capacity obtained by the use of a wet spirometer which had been modified to require a minimal amount of oral pressure to activate the equipment. The vital capacity measurements were made with the nostrils open and occluded. Ratios were computed between the two measures in the manner suggested by Kantner (7) to indicate efficiency of velopharyngeal closure. If the two measures were equal, a ratio of 1.00 was assigned. If the measure with the nostrils occluded exceeded that with the nostrils open, a ratio of less than 1.00 was assigned.

Another set of pressure scores was obtained by using an oral manometer designed to measure air pressure in ounces per square inch. This equipment was available for the testing of only 41 of the 103 subjects. Measures were taken with the nostrils open and also with the nostrils occluded and ratios were computed in the same manner as for the vital capacity measurements.

TABLE 1. Comparison of oral breath pressure ratios of a group of cleft palate subjects who achieved velopharyngeal closure and a group who did not achieve velopharyngeal closure as revealed by lateral head x rays.

| | *Closure Group* | | *No Closure Group* | | | |
|---|---|---|---|---|---|---|
| | *N* | *Mean Ratio* | *N* | *Mean Ratio* | *t* | $t_{.01}$ |
| Wet Spirometer | 34 | .99 | 40 | .68 | 5.98 | 2.36 |
| Manometer | 10 | .98 | 19 | .70 | 3.13 | 2.47 |

X-ray data were obtained for each subject phonating the vowel [u] and also sustaining the fricative sound [s]. Measurements were made of the degree of velopharyngeal opening under both of these conditions only for those subjects who had had cleft lips only or cleft palates that had been surgically repaired. The measurements were made in millimeters and represented the minimal distance between the palate and the posterior pharyngeal wall or lymphoid tissue.

## *Results*

A gross dichotomy was made between the subjects who achieved closure while sustaining either [u] or [s] and those who did not. The mean ratios for the wet spirometer and for manometer measurements for the two groups (closure and no closure) and results of the *t* tests are reported in Table 1. In spite of the fact that the degree of velopharyngeal opening was not taken into account for the group that did not achieve closure, the difference between the mean ratios for the two groups is significant beyond the 1% level.

The absolute oral breath pressure data also are of interest. They were obtained by use of the oral manometer and were available for 10 subjects with closure and 19 subjects with no closure. The subjects with closure did as well with the nostrils open as they did with the nostrils occluded, the means being 15.6 oz and 15.3 oz, respectively. In contrast, the no-closure group achieved a mean of 7.2 oz with the nostrils open and 11 oz with the nostrils occluded. The most reasonable explanation for the difference between the mean pressures achieved by the closure and no-closure groups with the nostrils occluded is that this difference is due to sampling and experimental error rather than to an inherent difference in the blowing ability of the two groups. Finally, the two means, each above 15 oz per square inch, obtained for the good closure group, place this group in the 'average' or 'moderately high' categories of intraoral breath pressure levels attained by normal children in Goddard's (*4*) study. On the other hand, the mean of 7.2 oz per square inch obtained for the poor closure group with the nostrils open falls in the 'low' pressure category established by Goddard for normal children. In fact, only 6% of Goddard's 200 normal subjects obtained scores of 7 oz per square inch or less and those that did were, with one exception, five-year-olds, her youngest age group. These data and clinical impressions resulting from the testing of a wide variety of subjects with clefts not included in the study suggest that air pressure of 8 oz per

TABLE 2. Correlation coefficients between various oral-breath-pressure, x-ray, and articulation measures obtained from a group of subjects with cleft palates.

| *Variables Correlated* | *N* | *r* | $r_{.01}$ |
|---|---|---|---|
| Wet spirometer ratio and manometer ratio | 40 | .65 | .39 |
| Wet spirometer ratio and Templin ratio | 99 | .40 | .27 |
| Manometer ratio and Templin ratio | 41 | .50 | .39 |
| Wet spirometer ratio and frequency of correct production of selected plosive sounds | 69 | .53 | .33 |
| Wet spirometer ratio and frequency of correct production of selected fricative sounds | 69 | .48 | .33 |
| Wet spirometer ratio and velopharyngeal opening while phonating [u] | 65 | −.69 | .33 |
| Wet spirometer ratio and velopharyngeal opening while producing [s] | 64 | −.62 | .33 |
| Manometer ratio and velopharyngeal opening while phonating [u] | 23 | −.58 | .51 |
| Manometer ratio and velopharyngeal opening while producing [s] | 23 | −.69 | .51 |
| Templin ratio and velopharyngeal opening while phonating [u] | 67 | −.45 | .33 |
| Templin ratio and velopharyngeal opening while producing [s] | 67 | −.51 | .33 |

square inch is the minimum amount necessary for the satisfactory production of consonant sounds requiring oral breath pressure.

As indicated earlier, eight subjects had cleft lips with no involvement of the palate and seven subjects had unrepaired clefts of the palate. These two groups probably represented the extremes of velopharyngeal competency. The mean wet spirometer ratio was 1.00 for the cleft-lip-only group and .36 for the unrepaired-palatal-cleft group. The mean Templin ratio was 1.00 for the cleft-lip-only group and .60 for the unrepaired-palatal-cleft group. Again the discrepancy in the means for the two groups emphasizes the importance of velopharyngeal competency both in developing oral pressure and in adequate articulation skills.

The strength of relationship for a number of pairs of variables was estimated by Pearson's correlation procedure. The obtained *r*s, reported in Table 2, are all significant beyond the 1% level. The relationship between the wet spirometer ratios and the manometer ratios is fairly substantial ($r = .65$). Inspection of the data, however, reveals that for some individuals one of the measures was a better predictor of velopharyngeal closure and articulation ability whereas for other individuals the other measure was a better predictor. For example, a few individuals who were tested on the manometer were able to impound air in the buccal cavity by valving between the tongue and palate since the manometer functions as a closed air-system. In these instances, the pressure measurements obtained on the manometer were spuriously high. However, these same individuals could not use this valving technique when being measured on the wet spirometer because it functions as an open air-system. As a consequence, discrepancies between the two pressure

ratios were occasionally observed for some individuals.

The *r* for estimation of the relationship between the wet spirometer ratios and the Templin ratios is .40. The rather low strength of this relationship is not surprising in view of the fact that production of many of the sounds tested in the Templin test require little, if any, oral breath pressure. Furthermore, factors other than inadequate pressure can account for the misarticulations.

An *r* of .50 was obtained by correlating the manometer ratios to the Templin ratios. Although this coefficient is somewhat higher than that reported for the wet spirometer ratios and the Templin ratios, it is based on fewer subjects. Since the reduction in the number of subjects extends the confidence interval, the difference between the two coefficients cannot be meaningfully interpreted. The two *r*s are not independent and no test of the difference between them is available.

The wet spirometer ratios were correlated with the frequency of correct productions of 36 plosive sounds and of 36 fricative sounds elicited in the Templin test. To minimize the effects of immaturity of articulation development as a basis for the misarticulations, only data pertaining to those subjects who were eight years of age or older were used. Although for this group the fricative sounds were generally more defective than the plosive sounds, the frequency of fricative errors did not correlate more highly with the wet spirometer ratios than did the number of plosive errors. The Subtelny and Subtelny data (*13*) reveal that a very small velopharyngeal opening is likely to result in defective fricatives, suggesting that the relationship between frequency of fricative misarticulations and oral breath pressure is not linear. If the relationship is not linear in this instance, it follows that the correlation coefficient will be reduced.

The elimination of the younger subjects reduced the number of manometer readings to the point where correlations between the frequency of correct productions of the 36 fricative sounds and of 36 plosive sounds and the manometer ratios were not justified.

The wet spirometer and manometer ratios were correlated with the extent of velopharyngeal opening as determined from the x-ray film. The negative correlations shown in Table 2 indicate that, in general, the lower the pressure ratios, the greater the velopharyngeal opening. These data do not reveal any important advantage in the use of x-ray films taken during the production of the vowel [u] or the continuant [s] in predicting adequacy of oral breath pressure as measured either by the wet spirometer or the oral manometer.

Two more correlations were obtained from the Templin ratios and the x-ray measurements. The degree of relationship between the articulation and x-ray measures is roughly comparable to that between the articulation and pressure measures. The correlation coefficient of —.51 obtained by relating the Templin ratio to the minimal velopharyngeal distance during the production of [s] is of a somewhat lower order than the coefficient of .78 obtained by Hagerty and Hoffmeister (*5*) when they studied the relationship between the minimal velopharyngeal distance during the production of [s] and an articulation score. The difference

may be due to the relative number of pressure and nonpressure sounds contained in the two articulation tests although Hagerty and Hoffmeister do not describe their articulation test in detail. The correlation coefficient of —.45 obtained from the Templin ratio and the minimal velopharyngeal distance during the phonation of [u] may be compared to Counihan's (*3*) results. He also obtained x-ray measurements of velopharyngeal closure while his subjects phonated the vowel [u]. These measures were correlated with articulation scores and scale values obtained from ratings of 'understandability.' The obtained coefficients were .33 and .41, respectively. Wildman reports somewhat higher correlations between various measures of velopharyngeal adequacy obtained from x rays during the production of [u] and the amount of nasal emission during the production of selected plosive and fricative nonsense syllables.

The role of oral breath pressure in the articulation of speech sounds can be studied by examining the results of the pressure measurements and articulation tests of those subjects whose scores on these tests were normal or nearly so. Data are from 54 subjects with wet spirometer ratios of .95 or better and 38 subjects with Templin ratios of .90 or better. The mean wet spirometer ratio of the 38 subjects with Templin ratios of .90 or better was .95 and of the total group of 103 was .79. On the other hand, the group of 54 with wet spirometer ratios of .95 or better obtained a mean Templin ratio of .88 while the total group of 103 obtained a mean Templin ratio of .75. High Templin ratios are thus more likely to be accompanied quite consistently by high wet spirometer ratios than the reverse. These data seem to substantiate the clinical impression that good articulation is almost always associated with a high pressure ratio whereas a high pressure ratio is not always associated with good articulation.

TABLE 3. Comparison of wet-spirometer and Templin ratios of subjects (N=14) with clefts of of the palate only and subjects (N=52) with clefts of the lip and palate.

| | *Lip and Palate* | *Palate Only* | |
|---|---|---|---|
| | *Mean Ratio* | *Mean Ratio* | *t* |
| Wet Spirometer | .84 | .64 | 2.32* |
| Templin Test | .80 | .65 | 2.82† |

*Significant at 5% level.
†Significant at 1% level.

For analysis of the articulation and pressure differences between the group of subjects who had a cleft of the palate only and those who had a cleft of both the lip and palate, only the subjects who had had surgical management of the palatal cleft were used. Comparisons were made between the wet spirometer ratios and the Templin ratios of the two groups. Table 3 reveals that the group with cleft palate only had significantly lower means than the group with cleft lip and palate. This finding is in agreement with that of Bzoch (*2*), who found that preschool children with clefts of the palate only had significantly more articulation errors than those with clefts of the lip and palate. Clinical reports have indicated that the nature and size of the cleft and the amount of tissue available for the group with cleft palates only are such that the results of physical management are frequently poorer

for this group than for the group with cleft lips and palates. It appears that there may be physical and physiological bases for the differences observed between the subjects with different types of clefts.

### *Discussion*

The correlations between the various measures used in this study are relatively low. The absence of higher relationships is not surprising in view of individual variability and the complexity of the speech process. It should be remembered that each of the measures was obtained only once for each subject. No research has been done to determine how stable the x-ray and pressure measures are on repeated tests of individuals with cleft palates. Furthermore, it is obvious that the adequacy of articulation development depends on more than the preciseness of velopharyngeal closure and the amount of oral breath pressure that the speaker can develop. Other possible factors, for example, are faulty learning and dental deviations. In fact, it is somewhat surprising to find that the relationships under study here are as strong as they are. Yet almost without exception subjects who were unable to develop adequate breath pressure showed marked deficiencies in articulation.

The cleft-lip-only group had normal air pressure and articulation skills comparable to those of physically normal children of the same age and sex. These findings suggest that the limitations imposed by the labial and dental deviations which typically accompany a cleft lip do not materially affect articulation proficiency, at least when they do not exist in conjunction with those structural deviations created by a cleft palate. On the other hand, the subjects in the unrepaired-palatal-cleft group did have appreciable articulation proficiency even though they had a marked lack of structural continuity and also major oral pressure deficiencies. The results provide support for the oftenmentioned observation that speakers are frequently able to compensate for oral structural deficiencies and that the adequacy of articulation skills cannot be predicted solely on the basis of the adequacy of the articulatory mechanism.

It would appear that prediction of articulation performance would be equally successful from either x-ray data or pressure measurements. However, the results do not suggest that one set of measures should be used to the exclusion of the other since the highest relationship between the x-ray data and pressure measurements is only .69. For some individuals x-ray data are more indicative of velopharyngeal competency whereas for others oral-breath-pressure measurements are better indicators.

The purpose for the employment of pressure measures is not to predict accurately what the articulation of a given subject will be but rather to assess better the bases for the articulation problems observed. Individual variability argues for many different kinds of observations. Pressure measurements in conjunction with x-ray and articulation data should be useful in identifying the etiological factors involved in speech problems and thus make possible more intelligent planning for future physical and therapeutic management.

*Summary*

Lateral head x rays, intraoral breath pressure measurements, and Templin articulation scores were obtained on 103 children with cleft lips and palates.

The main findings were as follows: (a) Measures of intraoral breath pressure were related to velopharyngeal competency and articulation scores. (b) All of the subjects with cleft lips only had, in general, normal oral pressure and articulation scores. (c) Subjects with good articulation scores almost invariably had high pressure scores; subjects with high pressure scores occasionally had relatively poor articulation scores. (d) Individuals with cleft lips and palates, in general, achieve better articulation scores and higher breath pressure than do individuals with cleft palates only. (e) Measurement of oral breath pressure of speakers with cleft lips and palates can be of major diagnostic value in establishing the bases for the articulation problems.

*References*

1. Black, J. W., The pressure component in the production of consonants. *J. Speech Hearing Dis.*, 15, 1950, 207-210.
2. Bzoch, K. R., An investigation of the speech of pre-school cleft palate children. Ph.D. dissertation, Northwestern Univ., 1956.
3. Counihan, D. R., A clinical study of the speech efficiency and structural adequacy of operated adolescent and adult cleft palate persons. Ph.D. dissertation, Northwestern Univ., 1956.
4. Goddard, Lisette, A normative study of oral pressure in children. M. A. thesis, Univ. Iowa, 1959.
5. Hagerty, R., and Hoffmeister, F. S., Velopharyngeal closure: an index of speech. *Plas. Reconstr. Surg.*, 13, 1954, 290-298.
6. Hudgins, C. V., and Stetson, R. H., Voicing of consonants by depression of larynx. *Arch. néerl. Phon. exp.*, 11, 1935, 1-28.
7. Kantner, C. E., The rationale of blowing exercises for patients with repaired cleft palates. *J. Speech Hearing Dis.*, 12, 1947, 281-286.
8. McWilliams, Betty Jane, Some factors in the intelligibility of cleft-palate speech. *J. Speech Hearing Dis.*, 19, 1954, 524-527.
9. Spriestersbach, D. C., Darley, F. L., and Rouse, Verna, Articulation of a group of children with cleft lips and palates. *J. Speech Hearing Dis.*, 21, 1956, 436-445.
10. Starr, C. B., The study of some characteristics of the speech and speech mechanisms of a group of cleft palate children. Ph.D. dissertation, Northwestern Univ., 1956.
11. Stetson, R. H., *Motor Phonetics—A Study of Speech Movements in Action.* (2nd ed.) Amsterdam: North-Holland Publ. Co. (for Oberlin College), 1951.
12. Stetson, R. H., and Hudgins, C. V., Functions of the breathing movements in the mechanism of speech. *Arch. néerl. Phon. exp.*, 5, 1930, 1-30.
13. Subtelny, Joanne D., and Subtelny, J. D., A study correlating consonant intelligibility with velopharyngeal and dental relationships in cleft palate speakers. Paper read at Amer. Speech Hearing Ass. Conv., 1957. (*J. Speech Hearing Res.*, 2, 1959, 353-360.)
14. Wildman, A. J., The role of the soft palate in cleft palate speech. *Angle Orthodont.*, 28, 1958, 79-86.

# Calling Attention to Stuttering

M. E. WINGATE

The phenomenon of stuttering has received concentrated attention in the past three decades and during this time many hypotheses have been advanced to explain its nature. Prominent among those which have developed in this period has been the 'semantogenic' or 'evaluational' theory of Wendell Johnson (*13, 15, 16*), which asserts that stuttering develops as a result of critical evaluation and disapproval of 'normal nonfluency' in the early stages of speech development and persists because the individual strives to avoid the nonfluency which was disapproved. This position thus makes disapproval a central factor in the etiology and persistence of stuttering. It stresses that one should not be critical of nonfluencies since disapproval has the effect of increasing the speaker's effort to avoid nonfluency, an effort which is reflected in exaggerated pausing, repetition, and strain. This admonition has also been stated inversely: it is bad practice to praise a stutterer for fluent speech since the net effect is to strengthen his conviction that he should never speak nonfluently (*17*, p. 458).

M. E. Wingate (Ph.D., University of Washington, 1956) is Assistant Professor of Speech, University of Washington. This article is based on a doctoral dissertation done at the University of Washington under the direction of Dr. Charles R. Strother and is an adaptation of a paper presented at the 1957 Convention of the American Speech and Hearing Association, Cincinnati.

Wischner (27), in his description of stuttering as a learned avoidant response, has based his thinking on Johnson's position. In answering the question of what the stutterer is attempting to avoid, Wischner quotes Johnson's statement that stuttering is an avoidance reaction 'designed to avoid the nonfluency which the individual has learned to fear and dread and expect' and 'an attempt to avoid the nonfluency that was originally disapproved.' Wischner agrees that the stutterer's anxiety is based on this original disapproval but adds:

> Within the present learning theory frame of reference, however, it would appear to be more accurate to say that the adult stutterer is avoiding, not nonfluency as such, but the original consequences (noxious stimulation) which attended the original nonfluent behavior (p. 329).

By 'the original consequences' Wischner means 'painful reactions' within the individual elicited by the reactions of those judging his speech.

The notion that being critical of a stutterer's speech only serves to make it worse has received widespread acceptance although support for it has been largely on a common-sense basis with only implicit or indirect support from a few experimental studies. Certain studies have reported relationships between the frequency of stuttering and (a) the stutterer's own focus on his speech (*14, 18, 19, 20, 23*); (b) the effect of various distractions (*1, 2, 19*);

or (c) varying degrees of situational stress (*8, 10, 23, 25*). Studies employing experimental conditions in which explicit attention is called to the stuttering in some way have been extremely few in number. Studies by Maddox (*22*) and Van Riper (*24*) were oriented in this direction: Maddox reported stuttering to increase when the subjects observed themselves in a mirror while speaking; Van Riper reported an increase in stuttering when the subjects were threatened with an electric shock for each instance of stuttering. These findings, especially those of the latter study, seem to have had the effect of making further work in this area appear unnecessary. More recently Sheehan (*24*), using the writings of Wischner as his point of departure, reported a study in which he had stutterers repeat each stuttered word until the word was said fluently. This procedure resulted in a decrease in stuttering. Sheehan views stuttering as an approach-avoidance conflict and interpreted these results as indicating an 'unlearning' of the 'avoidant response.' One could as well maintain that the obtained decrease in stuttering was occasioned by the permissive atmosphere of the experimental condition which effected a decrease in what is generally referred to in the literature as 'situational anxiety,' or that a 'set' to speak more fluently was communicated to the subjects.

In view of the theory that stuttering is aggravated if a speaker's speech difficulties are called stuttering and are labeled as undesirable, and because conclusive supporting evidence is lacking in the literature, further investigation would appear indicated to discover what happens to a stutterer's speech when his 'blocks' or moments of difficulty are evaluated directly and critically.

### *Problem*

The present study was designed to test the hypothesis that amount of stuttering will decrease under conditions in which emphasis is placed on speaking without stuttering and then calling attention to stuttering whenever it occurs.

### *Method and Procedure*

The subjects were 18 male stutterers ranging in age from 15 to 39, with a mean age of 22. All of the subjects had stuttered since childhood and all had had some amount of speech correction. Using Shulman's (*25*) criteria, subjects were eliminated who did not stutter on at least 2% of the words in the first passage spoken.

In this study experimental conditions were devised so that the subjects communicated verbally to another person, a procedure being considered a closer approximation of the real-life situation than would be the case in simply reading a paragraph, the technique used in many studies of stuttering. In this more 'communicative' type of situation, the subject gave directions to an assistant to do various simple tasks.

The material used consisted of three different passages, ranging in length from 251 to 257 words, made up of sentences averaging nine words in length, stating directions to do simple tasks. The three passages were presumably of equal difficulty; the experimental design, however, insured that any effects due to possible differences among them were eliminated.

TABLE 1. Reliability of judgments of words stuttered.

| *Judge* | *Correlation Between Judgments* |
|---|---|
| Experimenter (Original) | .85 |
| Experimenter (Playback) | .83 |
| Independent Observer | .87 |

Each subject performed under three different conditions. Condition I was a neutral condition designed to provide a measure of the stutterer's customary speech behavior against which to compare his performance in the experimental conditions. Therefore, Condition I was always given first and each subject served as his own control. Conditions II and III were counterbalanced in order of presentation after Condition I. At least 24 hours intervened between any two of the three conditions; for most subjects the interval was 48 hours. Previous studies (*8, 21, 27*) have shown that the effect of adaptation[1] dissipates rather quickly after an experimental session, 24 hours being more than adequate to eliminate this effect. Condition II required that the subject's communication be interrupted whenever he began to stutter and to accomplish this he had to speak into a microphone connected to earphones worn by the assistant. To equalize the effect of the apparatus itself, the microphone and earphones were used in all three conditions.

Each subject spoke the directions in each of the three passages to the assistant, who carried out each direction when the subject had finished speaking it. The subject spoke a different set of directions in each of the three conditions. In all three conditions the subject spoke the directions eight times in succession, in order to provide a measure of the adaptation effect. To eliminate any effects that might have resulted from differences between the passages, they were counterbalanced among the three conditions and each passage followed each of the other passages an equal number of times.

For all conditions the experimenter was stationed in an adjacent room behind a one-way mirror from where he could control the apparatus for the experimental conditions. Before assuming his station the experimenter instructed the subject in what he was to do in the current condition.

In Condition I he was to speak in his usual manner.

In Condition II he was to complete the passage quickly but would not be permitted to continue his communication until he had said a stuttered word fluently. Whenever the subject began to stutter, the connection between the subject and the assistant was broken and a small light and buzzer signalled

[1]In stuttering literature 'adaptation' refers to the progressive decrease in stuttering occurring with repeated reading of the same material.

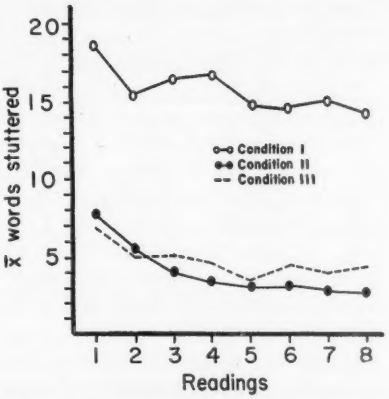



FIGURE 1. Frequency of words stuttered through successive readings in one neutral and two experimental conditions: Condition I, neutral, with no interruptions and no reminders; Condition II, experimental, with interruptions; Condition III, experimental, with reminders.

the subject that his communication was interrupted.

In Condition III he was told to speak as fluently as possible and that he would be reminded of each instance of stuttering, but could continue on. He was 'reminded' by the recording of each error on an electrically activated counter placed in front of him.

The experimenter followed the subject's course through each condition on copies of the particular passage in use, drawing a line through every word stuttered. A separate copy of the passage was used for each of the eight readings in any one condition. In all three conditions count was not made of those instances in which the subject checked himself when about to stutter but did not actually do so.

### *Reliability*

In order to check the experimenter's reliability in the judgment of words stuttered, a sample of each subject's performance in all conditions was recorded. These recordings were later played back, at which time the experimenter and another speech pathologist made independent judgments of the stuttering, marking them in the same manner as described above. The judgments from all recordings were pooled and three phi coefficients were computed to measure the correlations between the three sets of judgments (Table 1). The phi coefficient was considered the most appropriate statistic since it reflects the extent of agreement as to which words were stuttered rather than simple concurrence on the number of words stuttered. The correlations,

TABLE 2. Summary of $t$ test ($df$=17) and results for evaluation of differences in amount of stuttering between Condition I (neutral) and Condition II (interruption).

| *Reading* | $M_I$ | $M_{II}$ | $D_M$ | $\sigma_D$ | $t$ | $P$ |
|---|---|---|---|---|---|---|
| *1* | 18.72 | 7.78 | 10.94 | 2.64 | 4.14 | .01 |
| *2* | 15.61 | 5.67 | 9.94 | 3.22 | 3.09 | .01 |
| *3* | 16.72 | 4.94 | 11.78 | 3.92 | 3.01 | .01 |
| *4* | 17.01 | 3.78 | 13.23 | 4.23 | 3.06 | .01 |
| *5* | 14.83 | 3.61 | 11.22 | 4.09 | 2.74 | .02 |
| *6* | 14.50 | 3.72 | 10.78 | 3.58 | 3.01 | .01 |
| *7* | 15.17 | 3.28 | 11.89 | 4.31 | 2.76 | .02 |
| *8* | 14.56 | 3.28 | 11.28 | 3.96 | 2.85 | .02 |

TABLE 3. Summary of $t$ test ($df$=17) and results for evaluation of differences in amount of stuttering between Condition I (neutral) and Condition III (reminding).

| *Reading* | $M_I$ | $M_{III}$ | $D_M$ | $\sigma_D$ | $t$ | $P$ |
|---|---|---|---|---|---|---|
| *1* | 18.72 | 7.17 | 11.55 | 3.21 | 3.60 | .01 |
| *2* | 15.61 | 5.56 | 10.05 | 3.43 | 2.93 | .01 |
| *3* | 16.72 | 5.11 | 11.61 | 4.35 | 2.67 | .02 |
| *4* | 17.01 | 4.83 | 12.18 | 4.43 | 2.75 | .02 |
| *5* | 14.83 | 3.89 | 10.94 | 4.32 | 2.53 | .05 |
| *6* | 14.50 | 4.67 | 9.83 | 4.03 | 2.44 | .05 |
| *7* | 15.17 | 4.33 | 10.84 | 4.37 | 2.48 | .05 |
| *8* | 14.56 | 4.78 | 9.78 | 3.99 | 2.45 | .05 |

ranging from .83 to .87, indicate very satisfactory reliability of judgment and are consistent with figures reported in previous studies (*21, 24, 25*).

## *Results*

As indicated by the curves in Figure 1, there was a decreasing frequency of stuttering through the eight readings in all conditions. The curve for Condition I starts at a mean value of 18.72 stutterings in reading *1* and drops irregularly to a mean of 14.56 in reading *8*, a difference significant at the .10 level. This decrease is presumably due to the adaptation effect. The curves for Conditions II and III are quite similar. They start at means of 7.78 and 7.17, respectively, and drop in a fairly regular fashion to means of 3.28 and 4.78 in the last reading. These differences from first to last readings are significant at the .01 and .02 levels, respectively.

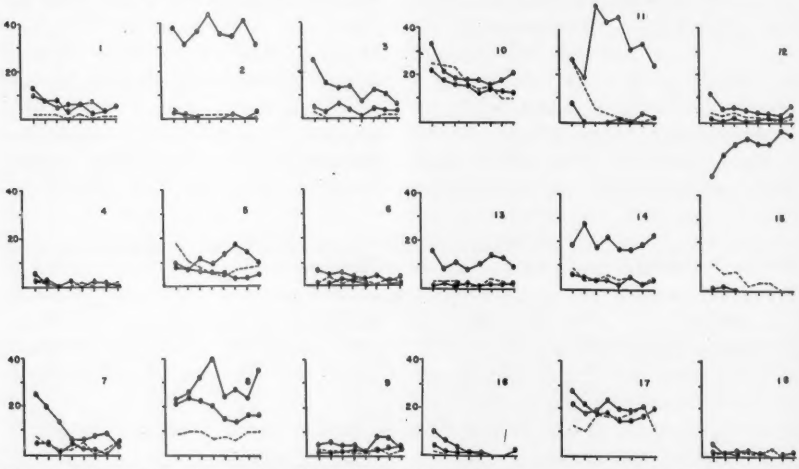


FIGURE 2. Frequency of stuttering of individual subjects. Lines connecting circles indicate Condition I, neutral, with no interruptions and no reminders; lines connecting dots indicate Condition II, experimental, with interruptions; dashed lines indicate Condition III, experimental, with reminders.

The effect of the two experimental conditions is clear; both differ significantly from the neutral condition in all of the eight readings (Tables 2 and 3). The differences between Condition I (neutral) and Condition II (interruption) are significant beyond the .02 level; the differences between Condition I and Condition III (reminding) are significant beyond the .05 level. Thus the null hypothesis tested can be rejected at these respective confidence levels. The differences between each of the experimental conditions and the neutral condition are evident from the first reading on, indicating that whatever factor operated to produce these differences was effective from the very beginning.

Analysis of the behavior of individual subjects yields some interesting results. Graphs plotted for each subject (Figure 2) reveal that all except four can be said to have shown some amount of improvement under one or both of the experimental conditions. These 14 divide themselves into two groups: eight who showed considerable or marked improvement, and six whose improvement was slight. In some measure the proportion of subjects showing only slight improvement is inflated due to the fact that roughly half of these subjects stuttered relatively little in the neutral condition, thus could show little improvement.

Inspection of the individual graphs reveals that, by and large, the curves for the two experimental conditions show less fluctuation and are flatter than those for the neutral condition. With very few exceptions these curves do not look like learning curves, that is, they do not show a pattern of decrement from the first to last reading which would depict learning. Analysis of the data for these curves lends support to this impression in that there is also no evidence of progressive improvement within any one reading; if each reading is divided into five 50-word segments and the number of stutterings in each of the five segments is tallied, these totals show little difference from one segment to another. This is true for the readings in all three conditions.

### *Discussion*

It is of considerable interest that the differences between the control and experimental conditions in this study are significant from the first reading. A very likely interpretation of this finding is that the subjects assumed a set to speak without stuttering (that is, a set to avoid speaking nonfluently), were able to maintain this set throughout the eight readings, and were able to improve their performance even though attention was called to their stuttering. The similarity of results for the two experimental conditions indicates that it made little difference whether attention was called to the block when it first appeared or after it was completed.

With the interpretation that the decrease in stuttering obtained in this study reflects the operation of a set to speak fluently, serious question is raised as to whether the stutterer's characteristic guiding motivation is simply to 'avoid nonfluency.' The results obtained suggest that if this were the stutterer's primary motivation he should be able to speak with less stuttering in ordinary situations also. On the contrary, then, it would appear that characteristically the stutterer is not so motivated and, in fact, is for some

reason unable to assume such a set. It is reasonable to believe that stutterers may avoid some social contacts, or avoid speaking in certain situations, or even avoid certain words;[2] and one should heed the stutterer's claim that he wishes to speak without stuttering. But to contend that he stutters in an effort to avoid nonfluency seems unwarranted, especially in view of the findings of the present study. Further, to explain the variation in appearance of stuttering in the individual as due to changes in 'fluency anxiety' begs the question. Undoubtedly anxiety is present, but it is questionable that this is simply fluency anxiety; it seems more plausible that the stutterer is not so much anxious about his speech as he is more or less anxious while speaking. The sources of this anxiety remain unidentified, and most likely vary with the individual.

It is felt that the findings of this study have important implications regarding the notion that stuttering reflects avoidance behavior, specifically avoidance of stuttering or avoidance of nonfluency. Emphasis on stuttering as avoidance behavior, as well as focus on the symptom per se, has drawn attention from the obvious fact that speaking is first and foremost man's basic means of communication. Speaking is not only activation of the speech mechanism resulting in certain sound patterns but is more importantly a process that involves (a) the conveying of a message and (b) a relationship of some kind between speaker and auditor. It would appear profitable to turn attention to these factors and give consideration to the place of stuttering in the total process of communicating and relating to another person.

This theoretical position suggests itself: the stutterer's motivation is to maintain communication, and this motivation is based on and expresses a deep need for interpersonal relations. Stuttering is learned and persists because, even though a block may be momentarily unpleasant when it occurs, it is nevertheless reinforced by virtue of the fact that it ultimately contributes to maintaining communication (and may have secondary gain value as well). That is, even when a stutterer blocks and his speech pattern is disturbed, he is still in a communicative relationship with his auditor. Assuming that the interpersonal relationship is important to the speaker, the block and the acts incident to overcoming it are as effective in maintaining this relationship as are the speech elements that precede and follow it. Perhaps the block is more effective to this end by virtue of possible attendant secondary gains (sympathy, attention).

*Implications for a Learning-Theory Interpretation of Stuttering.* In several studies note has been made of the similarity between the phenomena observed in studies of stuttering and the results obtained in laboratory studies of learn-

[2]No little attention has been accorded the apparent fact that the stutterer finds certain sounds and sound combinations difficult and thus tends to avoid them. Yet a large number of studies investigating the effect of formal elements of language structure on stuttering *(3, 4, 5, 6, 7, 9, 10, 11, 12)* have demonstrated that the mechanical features of speech are not directly related to stuttering. However, one of these studies, by Eisenson and Horowitz *(9)*, included an evaluation of the role of meaningfulness of the speech content and reported that this factor was directly related to stuttering. These authors made the significant comment that meaningfulness of speech content is a highly complex factor which interrelates with others of a personal nature.

ing. For example, the adaptation effect is likened to experimental extinction and the increase in stuttering following a time interval after adaptation is likened to spontaneous recovery. The results of the present experiment suggest that such changes may simply reflect fluctuations in the amount of anxiety present rather than the occurrence of learning.

Wischner (*27, 28, 29*) draws the parallel that stuttering resembles the instrumental avoidance type of learning. In this learning paradigm, a response is elicited to a conditioned stimulus or cue which has been associated with punishment. By making the appropriate response to the cue, the organism avoids the punishment. For Wischner, the cue is a word or a speech situation, the punishment is the painful reactions to disapproval, and stuttering is the avoidant response. However, this formulation of stuttering does not adhere closely to the learning theory model: the stuttering (avoidant response) continues to be disapproved (punished), that is, the stuttering does not avoid the punishment, yet stuttering persists. There are further difficulties in Wischner's position: although it emphasizes the unpleasantness which the stutterer anticipates, it does not account for the aspect of unpleasantness reported to occur during stuttering. Also according to his formulation, punishment is a necessary element in causing stuttering, a point criticized earlier.

A later publication by Wischner (*30*) lends support to an alternative hypothesis mentioned by him earlier (*27*), namely, the possibility that stuttering behavior is learned and persists because of the close association with tension-reduction when the word is finally completed. Such an hypothesis could well incorporate the factor of the discomfort during the stuttering act. It is possible that this neglected aspect of the stuttering sequence may offer the key to an adequate account of the mechanics of stuttering behavior.

If one wishes to conceive of stuttering as a form of learned behavior acquired through the agency of some painful stimulus source (fear, anxiety) it would seem that the instrumental-escape learning paradigm offers a better model. In this type of learning situation the animal learns to escape noxious stimulation by performing an act that stops the stimulation. The principle here is that by making a certain response the animal escapes from an aroused state of tension, and the response is reinforced because it results in tension reduction. Applied to stuttering, this model would offer the following account: in the course of speaking the stutterer reaches certain points wherein the level of tension in his attempt to communicate is raised. The speech musculature shares in this increased tension (probably in greater proportion since it is the muscle group in focus) and a block occurs. In the effort to continue communication, the individual strives to overcome the block, which effort is expressed as stuttering. Completion of the stuttered word brings tension reduction. The stuttering act is thereby instrumental in escape from aroused tension and is thus reinforced.

With this analysis in mind one might look again at the early period of speech development when stuttering is noted to arise most frequently. In acquiring the articulatory coordinations necessary for speech, the young child stumbles

and fumbles on many speech sounds as he attempts to express himself, and periods of fluency alternate with periods of nonfluency. If no additional stress is introduced into the performance of this activity then both types of behavior would be equally reinforced. Presumably due to increasing maturation, which is attended by greater facility in fine coordinative movements, articulation becomes increasingly precise and fluent speech thus receives more and more reinforcement. However, if additional stress is introduced into the speech situation, the child's level of tension in communicating is raised. Such additional stress may be created by any untoward emotional relationship between the child and his auditor(s), of which an adult's disapproving attitude toward the child's speech is but a part. Whatever the nature of the child's communicative behavior under these conditions, his speech, whether fluent or stuttered, receives more reinforcement because it is attended by a greater amount of tension reduction.

A conception of stuttering as instrumental-escape learning offers these advantages: (a) it is not embarrassed in its account of the role of disapproval in stuttering, since it need not formulate disapproval of speech pattern as a necessary nor the critical factor in stuttering, (b) it provides a theoretical position consistent with widely employed treatment procedures which aim to help the individual modify his stuttering responses and control his fears, and (c) it allows for a variety of dynamics underlying the tension surrounding the speech act. It also allows consideration of the still very real possibility that some stutterers may be inferior in terms of neuromotor function; that is, in some individuals more than others, fine coordinative acts (such as those involved in speech) might be more difficult to learn, might be more susceptible to the pressures of situational stresses, and blocks thus more readily occur.

*Summary*

After appraising the theory that stuttering is an attempt to avoid nonfluency and that conveying a critical attitude toward stuttering makes it worse, this study tested the hypothesis that stutterers would actually stutter less under conditions in which they were made aware of their stuttering. Recordings were made of the performance of 18 stutterers, speaking 250-word passages eight times in succession under three conditions, one in which they were permitted to speak in their usual manner and two in which attention was called to their stuttering. The results showed significantly less stuttering in the two experimental conditions. Analysis of the data suggests the interpretation that these results were due to subjects' assuming a set not to stutter, that is, to avoid speaking nonfluently; there was no evidence that learning occurred. The suggestion is made that consideration of the stutterer's motivation should take account of the economy of stuttering in the total process of communication. Note is made of implications in this study regarding description of stuttering as instrumental-avoidance learning. It is suggested that the instrumental-escape learning paradigm appears to provide a more suitable model than the instrumental-avoidance paradigm for analyzing stuttering in terms of modern learning theory.

## *References*


1. Barber, Virginia, Studies in the psychology of stuttering: XV, Chorus reading as a distraction in stuttering. *J. Speech Dis.*, 4, 1939, 371-383.
2. Barber, Virginia, Studies in the psychology of stuttering: XVI, Rhythm as a distraction in stuttering. *J. Speech Dis.*, 5, 1940, 29-42.
3. Brown, S. F., The influence of grammatical function on the incidence of stuttering. *J. Speech Dis.*, 2, 1937, 207-215.
4. Brown, S. F., Stuttering with relation to word accent and word position. *J. abnorm. (soc.) Psychol.*, 33, 1938, 112-120.
5. Brown, S. F., Theoretical importance of certain factors influencing the incidence of stuttering. *J. Speech Dis.*, 3, 1938, 223-230.
6. Brown, S. F., The loci of stutterings in the speech sequence. *J. Speech Dis.*, 10, 1945, 181-192.
7. Brown, S. F., and Moren, Adelaide, The frequency of stuttering in relation to word length during oral reading. *J. Speech Dis.*, 7, 1942, 153-159.
8. Dixon, C. C., The amount and rate of adaptation of stuttering in different oral reading situations. M. A. thesis, Univ. Iowa, 1947.
9. Eisenson, J., and Horowitz, Esther, The influence of propositionality on stuttering. *J. Speech Dis.*, 10, 1945, 193-197.
10. Hahn, E. F., A study of the relationship between the social complexity of the oral reading situation and the severity of stuttering. *J. Speech Dis.*, 5, 1940, 5-14.
11. Hahn, E. F., A study of the relationship between stuttering occurrence and phonetic factors in oral reading. *J. Speech Dis.*, 7, 1942, 143-151.
12. Hahn, E. F., A study of the relationship between stuttering occurrence and grammatical factors in oral reading. *J. Speech Dis.*, 7, 1942, 329-335.
13. Johnson, W., An open letter to the mother of a stuttering child. *You and Your Child*, April, 1941.
14. Johnson, W., A study of the onset and development of stuttering. In W. Johnson (Ed.), *Stuttering in Children and Adults*. Minneapolis: Univ. Minnesota Press, 1955.
15. Johnson, W., *People in Quandaries*. New York: Harper, 1946.
16. Johnson, W., Perceptual and evaluational factors in stuttering. In L. Travis (Ed.), *Handbook of Speech Pathology*. New York: Appleton-Century-Crofts, 1957.
17. Johnson, W., The Indians have no word for it: II, Stuttering in adults. *Quart. J. Speech*, 30, 1944, 456-465.
18. Johnson, W., and Rosen, L., Studies in the psychology of stuttering: VII, Effect of certain changes in speech pattern upon frequency of stuttering. *J. Speech Dis.*, 2, 1937, 105-109.
19. Johnson. W., and Sinn, Arlien, Studies in the psychology of stuttering: V, Frequency of stuttering with expectation of stuttering controlled. *J. Speech Dis.*, 2, 1937, 98-100.
20. Johnson, W., and Solomon, Anita, Studies in the psychology of stuttering: IV, A quantitative study of expectation of stuttering as a process involving a low degree of consciousness. *J. Speech Dis.*, 2, 1937, 95-97.
21. Jones, E. L., An investigation of stuttering in the light of a learned behavior, with special reference to experimental extinction and spontaneous recovery. Ph. D. dissertation, Univ. Iowa, 1950.
22. Maddox, J., Studies in the psychology of stuttering: VIII, The role of visual cues in the precipitation of moments of stuttering. *J. Speech Dis.*, 3, 1938, 90-94.
23. Porter, Harriet v. K., Studies in the psychology of stuttering: XIV, Stuttering phenomena in relation to size and personnel of audience. *J. Speech Dis.*, 4, 1939, 323-333.
24. Sheehan, J. G., The modification of stuttering through non-reinforcement. *J. abnorm. (soc.) Psychol.*, 46, 1951, 51-63.
25. Shulman, E., Factors influencing the variability of stuttering. In W. Johnson (Ed.), *Stuttering in Children and Adults*. Minneapolis: Univ. Minnesota Press, 1955.
26. Van Riper, C., The effect of penalty upon the frequency of stuttering spasms. *J. genet. Psychol.*, 50, 1937, 193-195.
27. Wischner, G. J., Stuttering behavior and learning: a program of research. Ph.D. dissertation, Univ. Iowa, 1947.
28. Wischner, G. J., Stuttering behavior and learning: a preliminary theoretical formulation. *J. Speech Hearing Dis.*, 15, 1950, 324-335.
29. Wischner, G. J., An experimental approach to expectancy and anxiety in stuttering behavior. *J. Speech Hearing Dis.*, 17, 1952, 139-154.
30. Wischner, G. J., Anxiety-reduction as reinforcement in maladaptive behavior: evidence in stutterers' representations of the moment of difficulty. *J. abnorm. (soc.) Psychol.*, 47, 1952, 566-571.

# Post-exposure Residual Effects of Low-Level Noise

WILLIAM J. TRITTIPOE

In a previous study (*2, 3*) the residual effects of low-level noise on the temporary threshold shift (TTS) were demonstrated. It was shown that noise of a level insufficient to produce a significant TTS will significantly modify the TTS following exposure to a more intense sound. Specifically, when a 15-minute exposure to white noise at 78 db SPL (which alone produced no TTS) preceded a three-minute exposure to white noise at 118 db SPL, the TTS following the second exposure was greater than when the same exposure was preceded by a control condition of 15 minutes of silence. This might be termed a 'pre-exposure residual effect.' The present study is concerned with the question of whether a post-exposure residual effect can be similarly demonstrated. That is, if a low-level noise which in itself is insufficient to produce a measurable threshold shift follows a high-level noise which does produce a measurable threshold shift, will the threshold shift be greater than when the high-level noise is followed by a period of silence of a duration equivalent to that of the low-level noise?

William J. Trittipoe (A.M., George Washington University, 1959) is Research Psychologist, Human Resources Research Office, George Washington University, Washington, D. C. This study was executed when he was Research Psychologist, Operational Applications Laboratory, Air Force, Cambridge Research Center, Bedford, Massachusetts.

## *Procedure*

*Terminology.* Presentation of sound at a level known to produce a TTS will be referred to as a 'principal exposure.' Presentation of sound at a level at which that sound alone produces no apparent TTS will be referred to as a 'sensitizing exposure.' When a sensitizing exposure precedes a principal exposure the former will be termed the 'sensitizing pre-exposure.' When the sensitizing exposure follows a principal exposure, the sensitizing exposure will be referred to as the 'sensitizing post-exposure.' If a period of silence precedes or follows the principal exposure, that period of silence will be termed, respectively, the 'silent pre-exposure' or the 'silent post-exposure.'

*Method.* The TTS was measured following two conditions of exposure to a principal exposure noise level: (a) a control condition (upper section Figure 1) in which a principal exposure was followed by a silent post-exposure; (b) an experimental condition (lower section, Figure 1) in which a principal exposure was followed by a sensitizing post-exposure. The noise level of the

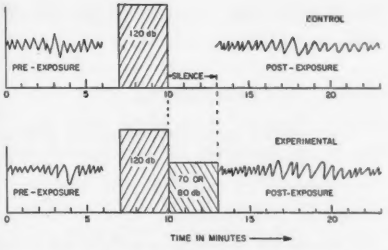



FIGURE 1. Schematic diagram of control and experimental conditions. The top section represents the control condition; the bottom section represents the experimental condition. The two conditions are identical, except that for the control a period of silence followed the 120-db noise exposure, and for the experimental condition a noise of low level followed the 120-db exposure

principal exposure was 120 db SPL. The duration of the entire test was 23 minutes. The first seven minutes were devoted to threshold measurements at 4000 and 6000 cps, alternately. This was followed by a three-minute principal exposure. Immediately after the principal exposure, there was a three-minute silent post-exposure for the control condition, a three-minute sensitizing post-exposure for the experimental condition. The post-exposure noise levels were 70 and 80 db SPL.[1] Following the sensitizing post-exposure, and following the silent post-exposure, recovery threshold measurements were made at 4000 and 6000 cps for 10 minutes.

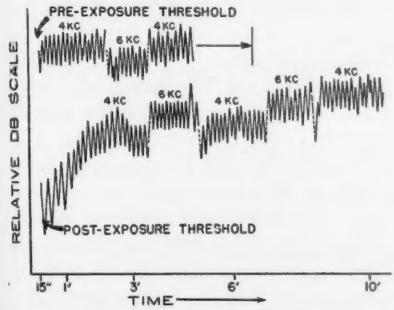



FIGURE 2. Measurement schedule for the test frequencies employed. Time is represented on the abscissa, and a relative db scale is represented on the ordinate. The test frequencies were 4000 and 6000 cps. The tracings on the upper half of the figure show the measurements made prior to sound exposure. The tracings on the lower half show the measurements made after sound exposure.

Figure 2 illustrates the threshold measurement schedule employed in the tests. Time is represented on the abscissa, and a relative decibel scale is represented on the ordinate. The tracings on the upper half of this figure represent the threshold measurements made prior to exposure to the principal exposure. The tracings on the lower half of the figure represent the sequence of measurements after the post-exposure condition. The difference in decibels between the measurements made before the principal exposure, and after the sensitizing post-exposure condition, defines the TTS. Only the latter of the two sets of measurements at 4000 cps, obtained prior to the principal exposure, was employed in the computations.

*Subjects.* Seven young adults with normal hearing served as listeners. Each listener contributed both ears to the tests. Each ear of each listener was exposed twice to the control condition, and twice to each experimental condition. (Each ear of two additional listeners was exposed to one run under

[1]The noise intensities are in decibels re 0.0002 dynes/cm$^2$. These sound pressure levels were measured by a standard 6-cc coupler and a 640 AA, Western Electric condenser-microphone system. The noise had a uniform spectrum level, and the noise levels reported represent the over-all output of the earphones from the coupler-microphone system.

each condition.) In addition, each ear of the nine listeners was exposed once to the higher sensitizing noise level (80 db) for six minutes. This was necessary to establish that the sensitizing noise levels by themselves produced no apparent TTS. These tests demonstrated 0 db of TTS after about 30 seconds following noise exposure. It was assumed that a 70-db exposure for six minutes also would have demonstrated 0 db of TTS.

The order of presentation of the experimental and control conditions was randomized, and at least 24 hours elapsed between successive exposures. If the listener's threshold had not recovered to within 5 db of his pre-exposure threshold, testing was delayed for at least another 24 hours.

*Instrumentation.* A Békésy-type, continuous-recording audiometer was used to obtain all threshold measurements. A single PDR-10 earphone, encased in a MX41/AR cushion and mounted in a binaural headset, was employed in all the tests. The earphones were not removed during any test including the period of silence for the control condition. The threshold test frequency was interrupted by an electronic switch. It was 'on' for approximately 0.4 sec and 'off' for 0.4 sec. The rise-fall time was 50 msec. All threshold measurements were made in an I.A.C. audiometric testing room. The ambient noise level in the testing room was 35 db SPL, with most of the energy concentrated below 70 cps. The noise exposures were of uniform spectrum with the band width limited only by the headphone's characteristics. For the principal exposure the noise was turned on slowly, that is, it took about seven seconds for the level to go from 0 db to 120 db SPL.

## *Results and Discussion*

The results averaged over the group are presented in Figure 3. The TTS is represented on the ordinate; the post-exposure measurement time is represented on the abscissa. Each point represents the average of 28 runs. An inspection of the curves in Figure 3 shows

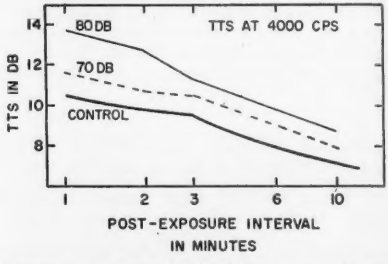



FIGURE 3. Average TTS (N=7) at 4000 cps as a function of the post-exposure measurement time. The thick, solid curve represents the control condition. The dotted curve and the thin, solid curve represent the 70- and 80-db post-exposure conditions, respectively.

that for the data averaged over the group, both the 70- and 80-db post-exposure conditions produced a greater TTS than the control condition. This differential between conditions is apparent even 10 minutes after sound cessation. Differences between the experimental and the control conditions were examined by $t$ tests for correlated measures. Significant differences were observed between the 80-db experimental condition and the control condition ($p<0.02$). However, the 70-db experimental condition and the control condition were not significantly different. Measurements made at 6000 cps for residual effects showed no significant

differences between the experimental and control conditions. In this analysis, each listener contributed one score under each condition (the TTS averaged over both ears for each listener for measurements made at 2-, 3-, and 6-min post-exposure time). The two listeners who failed to complete the tests were not included in this analysis.

All but one listener showed post-exposure residual effects. In only one case did the higher post-exposure noise level fail to generate a greater TTS than did the control. It should be noted that in a study concerned with pre-exposure residual effects (*2*) this same listener showed no residual effect.[2]

Harris (*1*) in an earlier study has shown that a 0 db TTS may not represent a complete recovery of the auditory system. By implication, the present study confirms Harris' conclusion and further demonstrates that under special conditions, low noise levels may have a residual effect on the hearing system.

It is concluded that the question asked in this study may be answered affirmatively; namely, that sensitizing post-exposure residual effects may be demonstrated.

[2]In a paper concerned with pre-exposure effects, Ward (*4*) has reported, using young adults as listeners, negative results. However, when he ran himself under the same conditions the results were positive. Dr. James M. Pickett has suggested that perhaps age is an important factor for demonstrating residual effects. While a precise statement cannot be made as to the age of the listeners in this study, it appears that the older listeners exhibit a relatively greater residual effect than do the younger.

## *Summary*

This study sought to answer the question of whether a sound, in itself insufficient to produce a temporary threshold shift (TTS), can modify the recovery threshold to a more intense sound. TTS measurements were obtained following two conditions of high-level noise exposure: (a) a control condition in which a period of silence followed a broad-band noise exposure; (b) an experimental condition in which the same exposure was followed by noise at levels which alone produce no measurable TTS. A 3-min exposure to a noise level of 120 db SPL was followed in the control condition by a 3-min exposure to silence and in the experimental condition by a 3-min exposure to noise of 70 and 80 db, noise which alone produced no measurable TTS. Statistical analysis demonstrated highly significant differences between the experimental condition (80 db) and the control. The initial question is interpreted affirmatively.

## *References*

1. Harris, J. D., On latent damage to the ear. *J. acoust. Soc. Amer.*, 27, 1955, 177-179.
2. Trittipoe, W. J., Residual effects at longer pre-exposure durations. *J. acoust. Soc. Amer.*, 31, 1959, 244.
3. Trittipoe, W. J., Residual effects of low noise levels on the temporary threshold shift. *J. acoust. Soc. Amer.*, 30, 1958, 1017-1019.
4. Ward, W. D., On "latent" auditory fatigue. *J. acoust. Soc. Amer.*, 31, 1959, 831-832.

# Factors Related to Lipreading

AUDREY ANN SIMMONS

Some hearing-impaired individuals learn to read lips very well while others do not acquire very much skill even after long periods of training. The present study is concerned with: (a) the measurement of the ability to lipread and (b) the factors in the individual that may be related to the ease or difficulty with which he reads lips. It is not concerned with various aspects of the educational procedures themselves.

## *Measurement of Lipreading*

One major obstacle to establishing a relation between different characteristics of an individual and his ability to read lips is the lack of a quantitative measure of lipreading ability. There have been various attempts to devise tests. In 1928 Day, Fusfeld, and Pintner (*1*) created a series of tests that were given in a face-to-face situation but the variability of speaking situations and among speakers contributed to low reliability.

The advent of motion pictures was accompanied by a series of developments in which both the test material and the speaker of the material could be held constant from trial to trial. In 1932 Mason (*11, 12*) described the use of films as an aid to instruction and later these same films were used as a test of comprehension. Heider and Heider (7) and Reid (*20*) also prepared films for test purposes.

An effort at standardization of a film test on a population of 761 people was made by Utley (*23, 24*). Utilizing the Utley test, Heider (*9*) found it unnecessarily long for children but she indicated that it was in general a good measuring device. DiCarlo and Kataja (*2*) found the Utley film too long even with adults but they also found a high correlation ($r = .77$) between the results of the Utley film and those of another film that had been designed for teaching. They also reported that hearing-impaired subjects gave lower scores than normal subjects not trained in lipreading, and this result was partially corroborated by O'Neill's (*15*) demonstration that normally hearing individuals utilize lipreading considerably in comprehending speech.

There appears to be no doubt that a filmed test of lipreading can test reliably some aspect of lipreading, but many therapists seem to feel that the ability tested by films is not the same as the one that they would like to test —the ability to read lips in the everyday face-to-face situation.

The present study represents an attempt to create a more valid measure than is now available of ordinary,

Audrey Ann Simmons (M.S., Washington University, 1945) is Director of Aural Rehabilitation and Instructor, Teachers Training College, Central Institute for the Deaf, St. Louis. This research has been partially supported by a grant (B-243) from the National Institute of Neurological Diseases and Blindness of the National Institutes of Health.

everyday lipreading through the use of an interview carried out in a face-to-face situation. Various scoring procedures, to be described below, yield a measure of lipreading derived from this interview situation, and this measure is used as an external criterion with which other tests can be correlated.

*The Interview Criterion.* Five judges, who served also as speakers, were selected to carry out and quantify the interviews. All five were therapists with two or more years of experience with hard-of-hearing adults.

Rating Scale. The judges were first asked to describe levels of ability in lipreading in such a way that a scale could be created. The following emerged as descriptive criteria under three general levels of lipreading ability:

*Good:* does not often miss what is said; lipreads without hesitation; follows shift in topic; can follow direct (to him) and indirect (between and among others) conversation; can lipread profiles and full face; can lipread variety of speakers; needs no special distinctness on part of speaker.

*Average:* sometimes hesitates and requires a filling in; eventually gets what is said either through repetition or rewording without an extreme breaking down of what was said; lipreads some people better than others; has some difficulty in following change in topic; has some difficulty with individual conversation; has some ease with speech if it is distinct; requires a simplification of language on part of speaker.

*Poor:* requires frequent repetition; requires written clues; has only limited range of topic; can't follow even when speech is reworded; can't pick up speech from context; needs to have lip movements exaggerated; needs slow rate; needs simple language.

Use of the Rating Scale. The five judges together engaged each of 24 hard-of-hearing subjects in conversation. The order in which the judges entered the conversation remained constant. Topics common to all interviews were: weather, television, news, sports, hearing aids, children, and community. The interview was carried out with speakers talking at a level just below the threshold level that was required by each patient to recognize some speech through hearing alone.

After each interview the judges reviewed the conversation together and then individually evaluated the ability of the subject on the scale that they had created. For this purpose a graphic evaluation was indicated on a scale laid out on a 10-cm line with points labeled *good, average,* and *poor,* marked at distances 2.5, 5.0, and 7.5 cm from the left end of the line. Numerical scores were thus obtained in centimeter units.

An analysis of variance of the judges' scores resulted in an estimate of the reliability coefficient of .92, a figure that indicates satisfactory reliability of the mean ratings used as criterion measures.

*Correlation Between Criterion and Filmed Tests.* Two silent movie tests were used, one by Utley (*23, 24*) and one by Mason (*11*). The Utley film ('How Well Can You Read Lips?') contains three kinds of material: Part I, a sentence test of 31 items including short phrases and full sentences; Part II, a word test of 36 words; and Part

TABLE 1. Scores on the tests of lipreading. (Numbers in parentheses on Utley test scores are values obtained by Utley.)

| | *Mean* | *S. D.* |
|---|---|---|
| Utley | | |
| Part I | 34.28 (33.63) | 13.34 (16.36) |
| Part II | 5.67 ( 6.85)* | 2.40 ( 3.80) |
| Part III | 5.58 ( 5.62) | 5.17 ( 4.48) |
| Total | 45.53 | 16.64 |
| Mason | 40.50 | 22.59 |
| Interview | 4.71 | 3.22 |

* Significant at the 5% level.

III, a comprehension test in which are shown six short scenes involving one or more speakers. Following each story in this third part is a series of five questions, after each of which the subject writes an answer of a sentence or word. The first two parts are in black and white, the stories are in color. The speaker in the first two parts is a young lady, while both males and females are involved in Part III. One point is given for each word correct in Parts I and II, and one point is given for each correct answer in Part III. The Mason film, in black and white, shows a young lady speaking 27 sentences. The subject's score is obtained by giving one point for each correct word in the sentences with a total possible score of 178 points.

TABLE 2. Intercorrelations between interview and filmed-test measures of lipreading.

| *Variable* | *Interview* | *Mason Film* | *Total Utley Score* |
|---|---|---|---|
| Part I Utley | .52* | .75* | .99* |
| Part II Utley | .37 | .78* | .68* |
| Part III Utley | .61* | .59* | .77* |
| Total Utley | .57* | .80* | |
| Mason Film | .61* | | |

* Significant at the 1% level.

*Results.* Table 1 shows the means and standard deviations of the measures for the 24 hard-of-hearing subjects obtained on each of the three tests used: the two films and the interviews. There was a significant difference ($t = 2.36$, significant at the 5% level) between results obtained by Utley in her study and the results of the present study on Part II of the Utley test. Except for this part (involving a word test) subjects in the present study on the average appeared to perform according to previous norms established for the Utley filmed test. No norms were available for the Mason film and, therefore, a similar comparison could not be made.

Intercorrelations among these measures of lipreading are shown in Table 2. There is a significant correlation between the results of the interview and the total score on the Utley and the Mason filmed tests. Both of these correlations, while not high, are significant at the 1% level. The filmed tests and the interview obviously do not measure exactly the same thing but these findings suggest that what they do measure involves a fair amount of ability that is common to the two kinds of tests.

All three parts of the Utley test are significantly correlated with the total score, and Part I has an extremely high weighting in the total score, which is necessarily the case, as may be noted from the relative sizes of the standard deviations reported in Table 1. The parts and the total score are also significantly correlated with the score on the Mason filmed test.

These intercorrelations may also be compared with previous reports of the reliability of these test instruments. Utley has reported a reliability coeffi-

cient for the sentence test (Part I) of .928, computed from the correlation between two forms of Part I. A similarly computed reliability coefficient for Part II is .797. The story test (Part III) apparently did not have two forms but is correlated with the other parts as follows: with Part I, .69; and with Part II, .56.

*Conclusions.* No attempt was made to compute a validity coefficient for the scores derived from the interviews. The assumption was that the way in which the scale was derived and the way in which the interview was carried out reflect a valid judgment of a person's lipreading ability. On this assumption the criterion of lipreading ability is defined as a score on a rating scale derived from the interview. The correlation between either filmed test and this criterion is a measure of the validity of the two filmed tests. While the correlations among parts of a filmed test or between two filmed tests are relatively high, this correlation between either one and the criterion is somewhat lower, although still significant. It appears, therefore, that these filmed tests do measure something related to lipreading ability but that the something is only a part of the general ability that can be scored in a face-to-face situation.

A somewhat different approach to this problem was reported by O'Neill and Stephens (*17*). This work, based upon the M.A. thesis (Ohio State University, 1956) of Stephens, reports somewhat lower correlations, significant at the 5% level, between parts of the Utley Test and the Mason Test. Correlations between a teacher's rating and the Utley Parts I and III were similar to the present results but again significant at the 5% level. There are several differences between the two studies. The 26 subjects used by O'Neill and Stephens were all taking formal lipreading instruction, whereas the subjects in the present study had had none. The O'Neill-Stephens group may have been more homogeneous because of this lipreading instruction, but no measures are reported that might yield such a fact. The rating in the O'Neill and Stephens study was made by only one teacher whereas the rating scores used here were based on the ratings of five judges, none of whom was the subject's teacher. Finally, O'Neill and Stephens use the low correlations between filmed tests and the single-teacher rating to question the validity of the filmed test and the validity of the rating. On the contrary, it is suggested here that the criterion of validity is defined in terms of a good rating scale, and that the validity of a filmed test can be ascertained only in terms of its correlation with such a criterion.

## *Factors Related to Lipreading*

There seems to be informal agreement among teachers of lipreading that their most successful students should have a high ability to synthesize meaningful wholes from fragmentary parts. This synthetic ability seems to be the only one on which there is relative unanimity of opinion. Much less agreement appears to be found among teachers with regard to other abilities.

*Previous Investigations.* Many more attempts to find factors that seem to be highly correlated with lipreading ability are to be found in the literature. These attempts have been focused on the individual's age, intelligence, edu-

cation, language and reading ability, synthetic ability, and the perception of color and form.

Age. Relating chronological age to lipreading ability, the Heiders (7) report a correlation coefficient of about .19 for children 8 to 17 years. Reid (*20*) reports the same number for students 10 to 22 years. These studies are supported by Utley's (*22*) conclusion that age is not an important variable in lipreading ability.

Intelligence. For a sample of children in advanced classes in schools for the deaf, Pintner (*18*) found no significant correlation between lipreading and intelligence, as measured by his Non-Language Test. Reid (*20*) did not find a correlation with intelligence in a sample of children from another school. On the other hand, O'Neill (*14*) found a correlation coefficient of .55 between lipreading and intelligence as measured by the Wechsler-Bellevue Performance Scale for a sample of normally-hearing young adults.

Education. Another negative finding of Reid (*20*) was the lack of correlation between lipreading and grade level or length of training at schools for the deaf. Utley (*22*) also found no correlation. But Heider (*9*) reported a correlation coefficient of .51 between the lipreading ability of school-age deaf children, as measured by the Utley filmed test, and the number of years in school. The Heiders (7) earlier had found a correlation coefficient of .54 between lipreading and educational achievement.

Language and Reading. Both language ability and ability for abstract thinking were claimed to be important for lipreading by Goldman (*4*) and by Utley (*22*), but O'Neill and Davidson (*16*) could find no correlation between reading ability and lipreading.

Synthetic Ability. Not only teachers' opinions but also some systematic results support the notion that synthetic ability is related to lipreading. With a tachistoscope, Kitson (*10*) determined a relation between synthetic and analytic perception and lipreading ability. A reading test and a completion test, presumed to be indicative of synthetic ability, yielded correlation coefficients of .67 and .65, respectively, with lipreading ability, in a group of 15 hearing-handicapped subjects. From the results of a similar experiment, Gopfert (*5*) concluded that synthetic ability should be dominant in the successful lipreader.

Lipreading Material. It appears that the perception of the phoneme plays a role in comprehension through lipreading, but investigators are not in agreement and some findings are in direct conflict. The Heiders (7) found a higher correlation between the ability to perceive vowels and lipreading than between the ability to perceive consonants and lipreading. On the other hand, O'Neill (*14*) reported that perception of the phoneme had the greatest effect on the identification of consonants and less on the recognition of vowels, words, and phrases. This difference might be attributed to the difference in population: the Heiders' group was deaf children; O'Neill's was hearing adults. Utley (*22*) concluded that stories served as better indices of performance than words or sentences. Nitchie (*13*) had said earlier that one of the most effective lipreading tests to determine the pupil's skill was story work. For Goldman (*4*), the knowl-

edge of the verb of the sentence seemed to be the most important factor.

Color and Form. Using a sorting task, the Heiders (*8*) found that good lipreaders sorted according to color while poor lipreaders preferred form. Using the Hanfmann-Kasanin Test, which requires the sorting of blocks according to height and size rather than form or color, O'Neill and Davidson (*16*) found a significant relation. They attributed it to the fact that the test measured nonverbal concept formation.

Other Factors. O'Neill (*14*) also explored personality adjustment as a factor relating to lipreading ability but found low and, in some instances, negative correlations. He found this to be true also when he related 'speech attitude' to lipreading performance.

Tactile reinforcement, the speech read by the eye and felt by the fingers, may be significant according to Gault (*3*). The Heiders (7) found performance in rhythm in dancing and in gymnastics related to lipreading ability ($r = .54$). Woolley (*26*) reported improvement in lipreading ability when employing the tachistoscope as a training device, but O'Neill and Davidson (*16*) reported no statistically significant relation between visual memory span and success in lipreading.

A survey of the literature yields a varied list of factors that might be related to lipreading but there is lack of agreement among investigators as to the degree of the relationship. There is also a wide variation in the population investigated. Some were normal hearing subjects, some congenitally deaf children, and others deafened adults. There is a lack of consistency in the measures of lipreading ability. The latter included speaking with subjects, using teacher's evaluations, and silent movie films. The fact that there are several such films available adds further to the possibility of differences. Synthetic ability and chronological age as significant factors relating to lipreading seem to be the only ones on which several investigators agree although their findings are inconclusive. Differences in correlation were found in intelligence, education, language, reading, perception of color and form, and visual memory. Factors such as personality, tactile reinforcement, and rhythm have not been investigated by more than one experimenter.

*Factors Used in the Present Study.* In searching for the most appropriate factors to be measured in the present study, leads from the past literature and from clinical experience, and also some mere guesses were used. General categories investigated, in addition to age and characteristics of the hearing, were: intelligence, reading, concept formation, synthetic ability, and certain perceptual abilities in vision and auditory rhythm. The specific tests used for these factors were as follows:

Intelligence. Form 1 of the Wechsler-Bellevue Scale (*25*) was used to measure intelligence. The complete battery consists of 10 tests, the first five of which are described as 'verbal,' the last five as 'nonverbal.' They are: Information, 25 questions for range of information; Comprehension, 12 questions on practical information and evaluation of past experience; Similarities, 12 pairs of nouns having a similar factor; Vocabulary, defining of 42 words; Picture Arrangement, six pictures for sequence story; Picture Completion, discovery of a missing part; Digit Sym-

bols, timed symbol-to-symbol identification; Block Design, timed colored-block assembly for designs. (Arithmetic and Object Assembly given to obtain a Full Scale Score were not used for correlation in the study and are not described here.)

Reading. The Iowa Silent Reading Test (*6*) consists of two somewhat diverse types of prose: Rate of Reading, in which subject reads for given time and is checked; and Comprehension, consisting of three subtests. These subtests are sentence meaning, 50 questions checking comprehension of sentence as a whole; paragraph comprehension, 12 paragraphs checking ability to generalize and identify details; and key words, 20 questions requiring generalization to a single word.

Concept Formation. The Hanfmann-Kasanin Test (*19*) is composed of 22 blocks of five different colors and four different shapes. These are to be sorted into categories. This test samples nonverbal concept formation.

Synthetic Ability. Three measures were utilized to test for synthetic ability. They are the Gestalt-Completion Test, 30 fragmented pictures of common objects projected upon a screen; Thurstone Mutilated Words, 26 fragmented printed one and two syllable words; Fragmentary Sentences, parts of 100 sentences requiring completion.

Visual Memory Span. Two tests were used to measure memory span. They are: Object Span Test, 21 slides of two or more common objects, as pictured by Dolch, for retention of object names; Digit Span Test, 43 slides containing eight 3-digit, nine 4-digit, six 6-digit, ten 5-digit, and seven 7-digit slides, for retention of digits presented visually.

Rhythm. Seashore Test (*21*) is a test designed to measure the sense of rhythm of an individual. The subject listens in rapid succession to two recorded rhythmic patterns, then indicates whether the patterns are the same or different. Form A is recommended for use with unselected groups, Form B for use where finer screening is desirable. One point is given each correct response.

*Procedure.* This study was designed to measure the possible relation between a wide variety of individual factors and lipreading ability. The subjects were 24 adults, 12 men and 12 women, who had been seen in the hearing clinics of Central Institute for the Deaf and for whom lipreading instruction had been recommended. None of the subjects had had any previous formalized lipreading instruction and it was assumed that since this instruction was needed, there would be a relatively high motivation for lipreading.

After an initial hearing test, each of the 24 subjects was interviewed by the group of judges in a manner described in the first section of this paper in connection with the establishing of a criterion for lipreading. Following the interview and the judges' rating of lipreading ability, one of the silent film tests was shown to the subjects as a group. No lipreading instruction was given during the time of the study. After all of the other factors had been measured, the group was shown the other filmed test of lipreading.

Each subject was seen twice by a psychometrist who gave, at the first visit, the Wechsler-Bellevue Test and, at the second, the Hanfmann-Kasanin Test.

TABLE 3. Hearing and age compared with lipreading measures. The last three columns give Pearson *rs* estimating relations between lipreading variables and the hearing and age variables.

| *Variable* | *Mean* | *S. D.* | *Pearson rs* | | |
|---|---|---|---|---|---|
| | | | *Interview* | *Mason* | *Utley* |
| Age | 47.13 | 9.64 | −.19 | −.25 | −.22 |
| Duration of Loss | 19.12 | 15.04 | .51* | .32 | .29 |
| Pure-tone Hearing Loss | 33.85 | 21.10 | .21 | .38 | .27 |
| Discrimination Score | 58.70 | 20.15 | −.26 | −.001 | −.36 |

* Significant at the 1% level.

The stimulus items of the Gestalt Completion, the Object Span, and the Digit Span Tests were flashed upon a screen that was viewed by all subjects as a group. At first, a nontest picture was shown so that the subjects could seat themselves for comfortable viewing and they all occupied the same seats in subsequent test sessions. A verbal signal, 'Ready, look,' preceded all test slides. The items of the Digit Span and the Object Span Tests were projected for 0.1 second. The Gestalt Completion Items were projected for 5 seconds.

The Iowa Reading Test, the Fragmentary Sentences, and the Mutilated Words were all paper-and-pencil tests and were administered to the subjects as a group. The standardized procedure for the Iowa Test was used, a time limit of 15 minutes was allowed for the Fragmentary Sentences, and five minutes were allowed for the Mutilated Words.

*Results.* The results will be ordered according to the factors studied.

Hearing and Age. The age and hearing characteristics of the sample

TABLE 4. Intelligence measures compared with lipreading measures. The last three columns give Pearson *rs* estimating relations between lipreading variables and intelligence variables.

| *Wechsler-Bellevue* | *Mean* | *S. D.* | *Pearson rs* | | |
|---|---|---|---|---|---|
| | | | *Interview* | *Mason* | *Utley* |
| Information | 11.7 | 2.7 | −.05 | .03 | .07 |
| Comprehension | 12.8 | 2.1 | .16 | .06 | .24 |
| Similarities | 12.9 | 2.6 | −.35 | −.24 | −.16 |
| Vocabulary | 12.1 | 2.6 | −.05 | −.03 | .07 |
| Verbal Score | 61.5 | 10.3 | .005 | −.09 | −.07 |
| Picture Arrangement | 9.9 | 3.1 | .18 | .46* | .48* |
| Digit Symbol | 10.4 | 3.4 | .42* | .42* | .50† |
| Block Design | 11.7 | 2.7 | .44* | .23 | .43* |
| Picture Comprehension | 12.4 | 2.3 | .08 | −.04 | .000 |
| Verbal IQ | 119.4 | 11.6 | −.008 | −.11 | −.08 |
| IQ Full Scale | 121.7 | 9.5 | .25 | .13 | .21 |

* Significant at the 5% level.
† Significant at the 1% level.

TABLE 5. Measures of reading skills compared with lipreading measures. The last three columns give Pearson *rs* estimating relations between lipreading variables and variables in reading skills.

| *Variable* | *Mean* | *S. D.* | *Pearson rs* | | |
|---|---|---|---|---|---|
| | | | *Interview* | *Mason* | *Utley* |
| Rate | 189.85 | 29.82 | .02 | .33 | .35 |
| Comprehension | 150.67 | 23.86 | .02 | .08 | .32 |
| Sentence Meaning | 173.63 | 22.81 | .02 | .33 | .44* |
| Paragraph Comp. | 148.83 | 24.78 | .03 | .31 | .38 |
| Key Words | 153.42 | 16.92 | .04 | .53† | .58† |

* Significant at the 5% level.
† Significant at the 1% level.

are shown in the first two columns of Table 3. The hearing loss is the average hearing loss at 500, 1000, and 2000 cps. With a mean hearing loss of only 33 db and an SD of 21 db, it is clear that a few of the subjects had very mild hearing losses, particularly for the frequencies that were averaged. These mild losses, however, were usually accompanied by sufficient discrimination loss that the entire group can be said to have been handicapped. The correlations between the four factors (age, duration of loss, amount of hearing loss, and discrimination score) and each of the three measures of lipreading ability are shown in the last three columns of Table 3. None of these factors is significantly correlated with the results of the filmed tests of lipreading and only the duration of loss is significantly correlated with the lipreading criterion that resulted from the interview rating.

Intelligence. Table 4 shows the relation between each of the subtests and subtotals of the Wechsler-Bellevue Test of intelligence and each of the three measures of lipreading ability. The bottom row shows the correlations between the full-scale IQ and lipreading ability. The lack of significant correlation between IQ and any measure of lipreading ability agrees with previous findings on normally hearing subjects (*16*) and on hard-of-hearing subjects (*18, 20*). The significant relation that was found between the Wechsler-Bellevue and the Mason filmed test by O'Neill (*15*) pertained to subjects with normal hearing.

Among the subtests, the Digit Symbol is significantly correlated with all three measures of lipreading, the Picture Arrangement Test with the two filmed tests of lipreading, and the Block Design Test with the Utley filmed test and with the interview criterion. Both the Digit Symbol and the Block Design Tests involve conceptualization of relations among parts without a verbal component. To be sure, picture comprehension and similarities are subtests that also require conceptualization, but these may be said to involve verbal concepts.

Reading Skills. The five scores available from the Iowa Reading Test are shown in Table 5, along with their correlations with the lipreading tests. The ability to extract sentence meaning and key words appears to be related to the ability that is tested by the filmed tests, but none of the factors in reading appears to be significantly

TABLE 6. Measures of synthetic ability compared with lipreading measures. The last three columns give Pearson *rs* estimating relationships between lipreading variables and variables in synthetic ability.

| *Variable* | *Mean* | *S. D.* | *Pearson rs* | | |
|---|---|---|---|---|---|
| | | | *Interview* | *Mason* | *Utley* |
| Gestalt Picture Completion | 15.29 | 6.93 | .38 | .25 | .29 |
| Mutilated Word | 94.17 | 4.75 | .32 | .28 | .24 |
| Fragmentary Sentence | 85.77 | 11.96 | .06 | .44* | .40* |

* Significant at the 5% level.

correlated with the lipreading ability that is measured by the interview rating. This finding is not in basic disagreement with previous work on reading.

Concept Formation. The Hanfmann-Kasanin Test is assumed to measure the ability to formulate concepts or to abstract. It seemed reasonable to suppose that the ability tested would be related to lipreading, particularly since the relation had already been shown to be significant by O'Neill and Davidson (*16*). In the present study the correlation between the Hanfmann-Kasanin and the Interview ($r = -.34$), Mason film ($r = -.36$) and Utley ($r = -.29$) all proved to be negative and not significant.

Synthetic Ability. Of the three tests that may be presumed to measure whatever is meant by the term synthetic ability (Table 6), the only one that appears to be significantly correlated with lipreading is the Fragmentary Sentences and here the correlation is significant only for the filmed tests. Insofar as these three tests may truly represent good measures of synthetic ability, one cannot conclude that synthetic ability is related to the general ability to lipread, as measured in the interview situation. There remains the possibility, of course, that synthetic ability is really necessary and that a good test of it is yet to be found.

Visual Memory Span. The two measures of visual memory span used were concerned only with the ability to recognize quickly and remember a group of objects or a line of digits. According to the results shown in Table 7, the Object Span Test was significantly correlated with all meas-

TABLE 7. Measures of visual memory span compared with lipreading measures. The last three columns give Pearson *rs* estimating relationships between lipreading variables with measures of visual memory span.

| *Variable* | *Mean* | *S. D.* | *Pearson rs* | | |
|---|---|---|---|---|---|
| | | | *Interview* | *Mason* | *Utley* |
| Object-Picture Span | 4.52 | .73 | .50* | .42† | .57* |
| Digit Span | 5.40 | .94 | .04 | .38 | .06 |

* Significant at the 1% level.
† Significant at the 5% level.

TABLE 8. Measures of rhythm compared with lipreading measures. The last three columns give Pearson *rs* estimating relationships between lipreading variables and rhythm-test results.

| *Test* | *Mean* | *S. D.* | *Pearson rs* | | |
|---|---|---|---|---|---|
| | | | *Interview* | *Mason* | *Utley* |
| Seashore A | 24.79 | 3.9 | .31 | .47* | .46* |
| Seashore B | 22.25 | 3.7 | .03 | .30 | .29 |

* Significant at the 5% level.

ures of lipreading and at least this aspect of visual perception appears to be an important factor. It is not at all clear why the visual span is related to lipreading when objects are used, but not when digits are used.

Rhythm. Assuming that the rythmic perception in the auditory system may be related to rhythmic perception in vision, it is reasonable to suppose that such ability would be important for capturing the rhythmic aspects of speech. The results in Table 8 show that the performance on the easier part of the Seashore Rhythm Test (Part A) is significantly correlated with the lipreading ability that is measured by the filmed tests, but not the lipreading ability measured in the interview. Another kind of rhythm, demonstrated in gymnastic tasks, yielded a correlation coefficient with lipreading ability of .54 according to the Heiders (7).

*Conclusions.* The correlations between the results on the various tests which sampled abilities in seven realms and lipreading ability are to be found in Tables 3 through 8. Table 9 is a summary of those test results that are significantly correlated with one or more of the measures of lipreading ability. From the information in this summary table the following conclusions may be drawn:

(a) A study of factors possibly related to lipreading ability must be lim-

TABLE 9. Relations between measures of lipreading ability and those factors found to be most highly correlated.

| *Factor* | *Test* | *Pearson rs* | | |
|---|---|---|---|---|
| | | *Interview* | *Mason* | *Utley* |
| Hearing | Duration of Loss | .51* | | .43† |
| Reading | Sentence Meaning | | | .44† |
| | Key Words | | .53* | .58* |
| Intelligence | Picture Arrangement | | .46† | .48† |
| | Digit Symbol | .42† | .42† | .50* |
| | Block Design | .44† | | .43† |
| Synthetic Ability | Fragmentary Sentences | | .44† | .40† |
| Visual Memory | Object-Picture Span | .50* | .42† | .57* |
| Rhythm | Seashore A | | .47† | .46† |

* Significant at the 1% level.
† Significant at the 5% level.

ited to the kind of lipreading ability measured. In the present study, only two test results are significantly related to both kinds of lipreading measures (filmed test and interview): the Digit Symbol subtest of the Wechsler-Bellevue Intelligence Scale, and the Object Picture Span, presumably one aspect of visual perception. These two tests apparently pertain to the ability to perceive rapidly and remember a number of items and the ability to associate symbols with symbols.

(b) Four factors appear to be related to the kind of lipreading ability that emerges from an interview rating. Two of them have been mentioned in (a) above. The other two are the duration of the hearing loss and the comprehension of relations as measured by the Block Design subtest of the Wechsler-Bellevue Intelligence Scale.

(c) Many of the factors appear to be related to the kind of lipreading ability that is measured by the filmed tests. Those factors that are significantly correlated with results of the filmed test appear to involve the ability to abstract meaning from printed material, the ability to perceive relations in a complex visual pattern, and the ability to reconstruct meaningful wholes from part stimuli that are impoverished in terms of spatial parts missing or insufficient time of presentation.

Conclusive interpretation of these results is difficult to make. It is reasonable to suppose that the longer a person has a hearing loss, the more he has come to depend upon information through lipreading, even though he has not had formal lipreading instruction. It is not certain whether other factors that are significantly related, such as the perception of spatial relations and the ability to perceive in spite of missing parts or quick presentation, constitute general abilities. If they do, the possibility remains that these are not genetically-determined f a c t o r s but rather the kind which can be improved with certain kinds of training. If one could, for example, improve a subject's object-picture span, then he might be a better candidate for lipreading instruction. On the other hand, a result from the Digit Symbol Test, presumably an aspect of biologically-determined intelligence, cannot easily be thought to be improvable.

*Summary*

An analysis was made correlating data on various tests with data on lipreading ability of 24 subjects with impaired hearing. The first purpose of this study was to develop a valid measure of lipreading ability through the use of a rating scale applied to the results of an interview, to represent an real-life lipreading situation. The ratings of the five judges were in very close agreement ($r = .92$) and they were significantly correlated with the scores obtained on the same subjects with two filmed tests of lipreading (Mason and Utley), but the correlations were not high.

Psychological measures of concept formation, intelligence, visual perception, reading, rhythmic perception, and synthetic ability, in addition to age, amount of hearing, and duration of hearing loss, were compared with measures of lipreading ability. Several abilities showed significant correlation with lipreading as measured with a filmed test, but only four specific test results were significantly correlated with lipreading as measured by rating from an interview.

## *Acknowledgment*

The author acknowledges with thanks the considerable help of Dr. Ira J. Hirsh and Dr. Robert C. Bilger of the Research Department of Central Institute for the Deaf.

## *References*

1. Day, H. E., Fusfeld, I. S., and Pintner, R., *A Survey of American Schools for the Deaf*. Washington: Nat. Res. Council, 1928.
2. DiCarlo, L. M., and Kataja, R., An analysis of the Utley Lipreading Test. *J. Speech Hearing Dis.*, 16, 1951, 226-240.
3. Gault, R. H., The interpretation of speech by tactual and visual impression. *Arch. Otolaryn.*, 4, 1926, 228-239.
4. Goldman, P., Untersuchungen über das ablesen vom munde bei taubstummen und spätertaubten. *Arch. ges. Phychol.*, 67, 1929, 441-504.
5. Gopfert, H., Psychologische untersuchungen über das ablesen vom munde bei ertaubten und horenden 2. *Kinderforsch*, 28, 1923, 315-367.
6. Greene, H. A., Jorgensen, A., and Kelley, V., *Iowa Silent Reading Tests*. Yonkers-on-Hudson: World Book Co., 1943.
7. Heider, F. K., and Heider, Grace M., An experimental investigation of lipreading. *Psychol. Monogr.*, 52, 1940, 124-153.
8. Heider, F. K., and Heider, Grace M., *Sixty-Ninth Annual Report of the Clarke School for the Deaf*. Northampton, Mass.: Clarke School Res. Div., 1936.
9. Heider, Grace M., The Utley Lip Reading Test. *Volta Rev.*, 49, 1947, 457-458, 488-490.
10. Kitson, H. D., Psychological tests for lipreading ability. *Volta Rev.*, 17, 1915, 471-476.
11. Mason, Marie K., A laboratory method of measuring visual hearing ability. *Volta Rev.*, 34, 1932, 510-516.
12. Mason, Marie K., A cinematographic technique for testing visual speech comprehension. *J. Speech Hearing Dis.*, 8, 1943, 271-278.
13. Nitchie, E. B., Tests for determining skill in lip-reading. *Volta Rev.*, 19, 1917, 222-223.
14. O'Neill, J., An exploratory investigation of lipreading ability among normal hearing students. *Speech Monogr.*, 18, 1951, 309-311.
15. O'Neill, J. J., Contributions of the visual components of oral symbols to speech comprehension. *J. Speech Hearing Dis.*, 19, 1954, 429-439.
16. O'Neill, J. J., and Davidson, JoAnn L., Relationship between lipreading ability and five psychological factors. *J. Speech Hearing Dis.*, 4, 1956, 478-481.
17. O'Neill, J. J., and Stephens, Mary C., Relationships among three filmed lipreading tests. *J. Speech Hearing Res.*, 2, 1959, 61-65.
18. Pintner, R., Speech and speech-reading tests for the deaf. *J. appl. Psychol.*, 13, 1929, 220-225.
19. Rapaport, D., *Diagnostic Psychological Testing*. Chicago: Year Book Pub., 1945.
20. Reid, Gladys, A preliminary investigation in the testing of lipreading achievement. *J. Speech Hearing Dis.*, 12, 1947, 77-82.
21. Seashore, C. E., Lewis, D., and Saetveit, J. G., *Manual of Instructions and Interpretations for the Seashore Measures of Musical Talents*. (rev. ed.) Camden: Educ. Dept., RCA, 1939.
22. Utley, Jean, Factors involved in the teaching and testing of lipreading ability through the use of motion pictures. *Volta Rev.*, 48, 1946, 657-659.
23. Utley, Jean, *How Well Can You Read Lips? Teacher's Lesson Manual and Motion Picture, Form A*. (accompanies film) Chicago: DeVry Corp.
24. Utley, Jean, A test of lipreading ability. *J. Speech Hearing Dis.*, 11, 1946, 109-116.
25. Wechsler, D., *The Measurement of Adult Intelligence*. (3rd ed.) Baltimore: Williams, Wilkins, 1944.
26. Woolley, Florence T., How we use the tachistoscope. *Hearing News*, 17, 1949, 3-4.

# Intelligibility and Associated Physiological Factors of Cleft Palate Speakers

JOANNE D. SUBTELNY

J. DANIEL SUBTELNY

Anatomic and physiologic studies of cleft palate subjects have yielded consistent evidence that the pathology in cleft palate speech results primarily from velopharyngeal inadequacy or dysfunction (*1, 3, 8*). The articulatory and quality characteristics of cleft palate speech have been discussed extensively in speech literature (*4, 5, 6, 9*). Despite these facts, there is a definite need to specify the physiologic parameters of cleft palate speech defects. This specification is important to an understanding of the influence of structural and functional deviations upon the intelligibility of speech sounds. Such an understanding would make it easier to determine which sounds can or cannot be achieved or improved. These considerations should determine to a large extent the ultimate objectives of treatment, the therapeutic emphasis, and the techniques to be employed.

The basic purpose of this study was to examine specific articulatory errors in cleft palate speech in association with measurements of the velopharyngeal and dental relationships. The specific aims of this study are as follows:

(a) To define, by intelligibility testing, articulatory errors in cleft palate speech.

(b) To determine statistically whether articulatory errors in cleft palate speech can be categorized phonetically.

(c) To study relationships between specific types of articulatory errors and velopharyngeal measurements.

(d) To study relationships between specific types of articulatory errors and measurements of dental relationships.

## *Procedure*

The study of possible relationships between articulatory errors in cleft palate speech and physiologic factors made necessary the attainment of high-

Joanne D. Subtelny, (Ph.D., Northwestern University, 1956) is Research Associate, Eastman Dental Dispensary, Rochester, New York, and J. Daniel Subtelny (D.D.S., University of Pennsylvania, 1947) is Head, Department of Orthodontia, Eastman Dental Dispensary, and Assistant Professor of Dentistry, School of Medicine and Dentistry, University of Rochester. This paper was presented in part at the 1957 Convention of the American Speech and Hearing Association, Cincinnati. The investigation was supported in part by research grant D-553 from the National Institute of Dental Research of the National Institute of Health, Public Health Service. Radiographic records were obtained at The Cleft Palate Center of the Chicago Professional Colleges, University of Illinois, Chicago.

fidelity tape recordings of an articulation test, x-ray registrations, and phonetic evaluations of the tape recordings.

*Subjects.* The subjects for this study were 27 cleft palate individuals ranging from 16 to 36 years of age. All subjects had surgically repaired clefts of the velum and hard palate. Nineteen subjects had complete unilateral clefts of the lip and palate, five had complete bilateral clefts of the lip and palate, and three had clefts of the hard and soft palates with no involvement of the alveolar process or lip.

Radiographic records and high-fidelity tape recordings of an articulation test were accumulated for these subjects.

*Articulation Tests.* The purpose of the articulation test was to determine whether or not subjects could produce identifiable consonants appearing in simplified and controlled phonetic environments. For this reason nonsense syllables, which do not provide contextual clues to the identity of the consonants, were used to test articulation.

The test material consisted of 48 nonsense syllables which included an equal number of consonant-vowel and vowel-consonant combinations. Consonants were selected to represent the three major classifications of consonants. Plosives [p] and [b], [t] and [d], [k] and [g]; fricatives [f] and [v], [s] and [z]; and glides [l] and [r] were combined with vowels [i], [u], [æ], and [ɑ]. Consonants were combined with each of the four vowels so that each consonant appeared twice in the initial and twice in the final position. For example, the consonant [p] appeared in syllables [pi], [pɑ], [æp], [up].

Each subject read the 48-syllable articulation test with five-second intervals between syllables. Tests were recorded by use of the Magnecorder, Model PT-6, and the Electro-Voice, Model 650 microphone. To minimize practice effect in listening, the order in which syllables were recorded was randomized for each subject.

The tapes were played back in a sound-treated room for seven speech correctionists who made phonetic transcriptions. These transcriptions were analyzed and the results averaged according to incidence of error and type of error.

*Physiological Factors.* To determine the physiological factors associated with articulatory errors, measurements were made from tracings of midsagittal cephalometric laminagraphs.[1] Films were taken under two conditions: with the teeth in occlusion and during phonation of [u]. Head stabilization was insured by a special head-holding device. Tape recordings of the vowel phonation, obtained simultaneously with the radiographic exposure, were later played back for phonetic identification of the vowel produced.

Two measurements were made from the films. The first, to quantify the degree of velopharyngeal constriction during sustained phonation of [u], was made at the point of closest approximation between the posterior aspect of the velum and the anterior margin of the posterior pharyngeal wall. The second was obtained from films taken with the teeth in occlusion and was used to define the anteroposterior rela-

[1]Laminagraphs are x rays obtained through the utilization of a body-sectioning radiographic technique. A specific lamina or layer of body structure can be visualized radiographically without the superimpositioning of structures on either side of the desired lamina.

tionships between the upper and lower incisors.

In normal dental relationships, the lingual surfaces of the upper incisors overlap or contact the labial surfaces of the lower incisors. A reading of 0 mm indicated that a normal anteroposterior relationship of the incisors existed. Positive readings indicated a degree of overjet or a protruded position of the upper (maxillary) incisors in relation to the lower (mandibular) incisors. Similarly, minus readings indicated a retruded position of the upper incisors in relation to the lower incisors. This method was utilized as a descriptive index of anterior dental relationships and does not offer reliable information pertaining to possible lateral constriction in the maxilla. A buccolingual constriction of the maxillary dental arch cannot be visualized in a midsagittal laminagraph. Despite this limitation, the evaluation of the relationship of the anterior teeth seems indicated. It has been found that in hypernasal cleft palate subjects, significant deviations in tongue posture were associated with severe retrusions of the maxilla (*8*). Maxillary retrusions would be manifested by a lingual positioning of the maxillary incisors relative to the mandibular incisors. For this reason, the relationship of the anterior teeth was measured.

*Results*

*Articulatory Errors.* The average number of articulatory errors on the 48-syllable test was found to be 11.03. For a better definition of the consonant types most commonly misarticulated, the average number of errors for each consonant pair was determined. These were as follows: [s] and [z], 2.79; [p] and [b], 2.17; [f] and [v], 1.80; [k] and [g], 1.72; [t] and [d], 1.42; and [l] and [r], 1.15. A decrease in the average number of misarticulations is indicated by the order reported, but only the mean for the pair [s] and [z] was significantly different from the means of the other pairs. When consonant pairs were grouped according to phonetic classifications, an obtained $t$ of 2.32 revealed that the incidence of fricative errors was significantly greater than the incidence of plosive errors.

Correlation coefficients (Table 1) were obtained to determine whether or not misarticulations on one pair of consonants were related significantly to misarticulations on other consonant pairs, and to determine whether or not a pattern of articulatory errors could be established. Such a pattern might be studied for some common physiologic

TABLE 1. Correlation coefficients estimating relationships between different types of articulatory errors of adult cleft palate speakers.

| | [f], [v] | [t], [d] | [s], [z] | [l], [r] | [k], [g] |
|---|---|---|---|---|---|
| [p], [b] | 0.20 | 0.32 | 0.18 | 0.17 | 0.62* |
| [f], [v] | | .49† | .22 | 0.42† | 0.44† |
| [t], [d] | | | .50* | 0.36 | 0.50* |
| [s], [z] | | | | 0.15 | 0.17 |
| [l], [r] | | | | | 0.20 |

* Significant beyond 1% level.
† Significant beyond 5% level.

factor associated with the specific articulatory defects in cleft palate speech. Consonant pairs were not, in general, found to be related to other consonant pairs on frequencies of errors. Exceptions to this were the significant correlations between plosives. As shown in Table 1, correlations significant beyond the 1% level were found between labial plosives and lingua-velar plosives; between lingua-alveolar plosives and lingua-velar plosives; between lingua-alveolar plosives and lingua-alveolar frictatives. Since these three groups of plosive consonants have differing positions of articulations, the suggestion that articulatory errors show similarity in position of articulation is not supported. However, the significant interrelationships indicate that there was some consistency of frequency of articulatory error during the production of sounds with abrupt onset characteristics.

No significant relationship was found between errors on [s] and [z] and errors on [f] and [v] even though these two fricative pairs are similar in manner of production, both being sustained sounds produced by forcing the breath stream through constricted areas within the vocal tract. It was also not found that glide errors were related to errors on any of the other consonant pairs, if the 1% level of confidence is used as a criterion.

*Physiologic Factors.* Dental and velopharyngeal measurements revealed that marked individual variation existed in both areas. Anterior dental relationships were found to vary from a slight overjet of +3.00 mm to severe maxillary retrusions of −11.00 mm. If significance can be attached to a mean of −2.26 mm, it can be stated that this sample of adult cleft palate subjects did not show severely disturbed anteroposterior relationships of the incisors.

Velopharyngeal measurements ranged from complete closure to a marked openness of 16 mm. The mean for velopharyngeal opening during phonation was 5.41 mm. Statistical correlation of velopharyngeal and dental measurements showed these two measurements were not related significantly.

*Total Articulatory Errors, Velopharyngeal Measurements.* On the basis of total number of articulatory errors, subjects with the highest degree of articulatory proficiency were those whose associated velopharyngeal relationships closely approximated the normal. Ten subjects were found to have relatively good articulation. In eight of these 10 subjects velopharyngeal opening did not exceed 1.5 mm. This finding tends to support the premise that articulatory proficiency is associated with near normal velopharyngeal function.

While a positive relationship between good velopharyngeal function and proficiency in total articulation was indicated, the results in regard to total articulation scores for the entire sample do not permit the conclusion that there is a well defined relationship between total articulatory proficiency and velopharyngeal function. The number of total errors, as determined by the nonsense-syllable test utilized, did not increase linearly as the degree of velopharyngeal opening increased. Some subjects were able to achieve far better total articulatory scores than others with essentially the same degree of velopharyngeal opening. In short, marked variation in total intelligibility scores for specified degrees of velopharyngeal opening was revealed.

TABLE 2. Correlation coefficients estimating relationships between specific articulatory errors and measurements of dental relationships, and between specific articulatory errors and measurements of velopharyngeal openings in adult cleft palate speakers.

| *Consonants Tested* | *Dental Measurements* | *Velopharyngeal Measurements* |
|---|---|---|
| | *r* | *r* |
| [p], [b] | −0.14 | 0.56* |
| [f], [v] | 0.03 | 0.36 |
| [t], [d] | −0.27 | 0.36 |
| [s], [z] | −0.06 | −0.03 |
| [l], [r] | 0.10 | 0.03 |
| [k], [g] | −0.31 | 0.69* |
| Dental Measurements | | −0.33 |

* Significant beyond 1% level.

*Specific Articulatory Errors, Velopharyngeal Measurements.* Since glides, plosives, and fricatives have essentially different methods of production, there seems a possibility that there may be also a difference in their intelligibility when produced by subjects with cleft palate aberrations. Considering this possibility and the clinical importance of determining which consonants are most consistently defective with associated velopharyngeal inadequacy, study of the correlations between specific articulatory errors and velopharyngeal measurements was indicated. These correlations are presented in Table 2.

As a group, plosive errors were found to be consistently related to velopharyngeal inadequacy. Correlations, significant beyond the 1% level, in spite of the restricted range for articulation scores, were found between errors on labial plosives and velopharyngeal openings (.56); and between errors on lingua-velar plosives and velopharyngeal openings (.69).

As indicated in Figure 1, both labial and velar errors increased appreciably

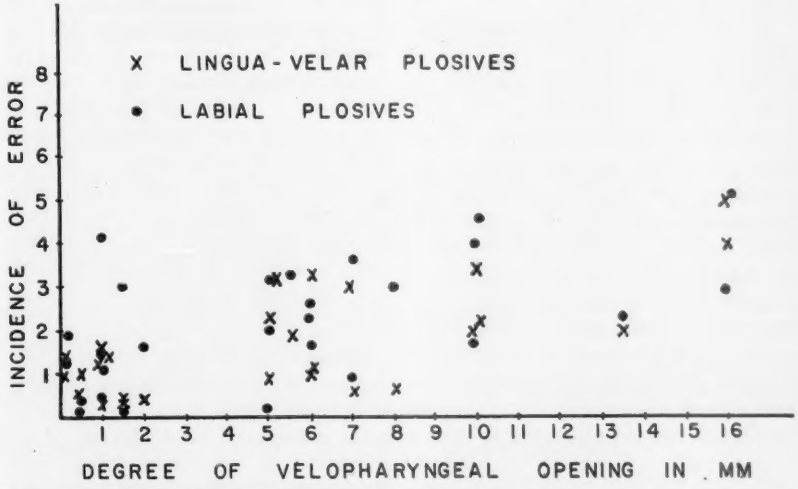



FIGURE 1. Incidence of lingua-velar and labial plosive errors with associated velopharyngeal measurements for each of 27 adult cleft palate speakers.

when the velopharyngeal opening equalled or exceeded 5 to 7 mm. Relatively few velar and labial errors were found when closure was not complete but within 2 mm of being complete. On the basis of these data, it would seem that complete velopharyngeal closure is not essential for production of plosives but that intelligible plosives frequently are not produced when velopharyngeal openings exceed 5 to 7 mm.

No significant relationships were found between fricative errors and degree of velopharyngeal openings. Fricative errors, abundant regardless of differences in velopharyngeal openings, were found to be almost twice as common as plosive errors in speech produced by subjects with small velopharyngeal openings (within 2 mm of complete closure). Cleft palate subjects, with essentially the same small degree of velopharyngeal opening, were better able to produce identifiable plosives than fricatives.

As shown in Table 2, no significant relationship was found between glide errors and velopharyngeal openings.

*Articulatory Errors and Dental Measurements.* The total number of articulatory errors was not found to be consistently or significantly related to measurements made between upper and lower incisors. To illustrate this, the five subjects with the highest number of articulatory errors were found to have dental relationships ranging from an overjet of +3.00 mm to a maxillary retrusion of −5.0 mm, and the five subjects with the fewest articulatory errors had dental relationships ranging from an overjet of +2.5 mm to a retrusion of −5.00 mm. No significant correlations were found between defectiveness in any of the consonant pairs studied and the degree of anterior malocclusion (Table 2).

## *Discussion*

Recently research attention has been focused upon organized and detailed investigations of the articulatory deficits of cleft palate children (*2*, *5*, 7) and adults (*3*, *4*, *9*). The incidence of articulatory errors reported in the present study is essentially in agreement with the results of other investigators who have studied a group of cleft palate speakers of comparable age (*4*, *9*). Specifically, this applies to the high incidence of sibilant defects and the comparatively low incidence of glide defects.

The present study of articulatory defects with accompanying physiologic data has yielded results which suggest that the types or the phonetic characteristics of articulatory errors in cleft palate speech are influenced by the degree of velopharyngeal opening. For example, the incidence of fricative errors was considerably higher than the incidence of plosive or glide errors in the speech of cleft palate subjects with small velopharyngeal openings (2 mm or less). Subjects with large velopharyngeal openings also had a high incidence of fricative errors but these subjects tended to have also a comparably high incidence of plosive errors.

Stated in other terms, present results suggest that a specific degree of velopharyngeal opening has an influence upon consonant intelligibility which may vary as a function of phonetic classification. At the present time, there seems little research evidence to sup-

port this suggestion. There is evidence, however, that velopharyngeal opening in cleft palate speakers has an influence upon nasality, or the degree of distortion, which varies as a function of the specific vowel produced (*6, 8, 9*). Because nasality has been found to vary as a function of the vowel, it has been stated that certain vowels have greater predictive value than others in estimating total nasality in connected speech (*8, 9*).

Present results, obtained from a sample limited to 27 cleft palate subjects, failed to reveal a direct relationship between velopharyngeal opening and total speech intelligibility. The marked variation in total intelligibility scores for specified degrees of velopharyngeal opening does not make it possible to predict total speech intelligibility from an associated measure of the degree of velopharyngeal opening. However, the significant correlations between plosive intelligibility and velopharyngeal opening suggest that the specific evaluation of plosive intelligibility has value in predicting velopharyngeal adequacy from the results of speech analyses.

Compositely, present results strongly indicate the need for continued study of correlation between speech performance and associated physiological factors of cleft palate speakers. Clinical implications derived from these data should be considered tentative until the results of a comprehensive study of a larger cleft palate sample are provided.

## *Summary*

Cephalometric laminagraphs and tape recordings of a nonsense-syllable articulation test were obtained from 27 adult cleft palate speakers. Laminagraphs were traced and analyzed to quantitate velopharyngeal and dental relationships. Tape recordings were played back to a panel of listeners for phonetic transcription. The results were analyzed and then correlated with the physiologic measurements.

Statistical analysis of consonant errors revealed the incidence of fricative errors was significantly higher than the incidence of plosive or glide errors. Intercorrelations of errors on plosive consonants were found to be significant. Correlations between measurements of velopharyngeal opening and plosive intelligibility also were found to be significant. On the basis of these findings, the specific evaluation of plosive intelligibility is suggested as useful in evaluating adequacy of velopharyngeal function. Since plosive errors were interrelated, irrespective of differing positions of articulation, the plosive manner of production is interpreted as a consistent attribute of defectiveness in cleft palate speech when an appreciable degree of velopharyngeal opening exists.

## *References*

1. Buck, M., Post-operative velo-pharyngeal movements in cleft palate cases. *J. Speech Hearing Dis.*, 19, 1954, 288-294.
2. Bzoch, K. R., An investigation of the speech of pre-school cleft palate children. Ph.D. dissertation, Northwestern Univ., 1956.
3. Counihan, D. T., A clinical study of the speech efficiency and structural adequacy of operated adolescent and adult cleft palate persons. Ph.D. dissertation, Northwestern Univ., 1956.
4. McWilliams, Betty Jane, Articulation problems of a group of cleft palate adults. *J. Speech Hearing Res.*, 1, 1958, 68-74.

5. Spriestersbach, D. C., Darley, F. L., and Rouse, Verna, Articulation of a group of children with cleft lips and palates. *J. Speech Hearing Dis.*, 21, 1956, 436-445.
6. Spriestersbach, D. C., and Powers, G. R., Nasality in isolated vowels and connected speech of cleft palate speakers. *J. Speech Hearing Res.*, 2, 1959, 40-45.
7. Starr, C. D., A study of some of the characteristics of the speech and speech mechanisms of a group of cleft palate children. Ph.D. dissertation, Northwestern Univ., 1956.
8. Subtelny, Joanne D., A laminagraphic study of nasalized vowels produced by cleft palate speakers. Ph.D. dissertation, Northwestern Univ., 1956.
9. Van Hattum, R. J., The interrelationships among measures of articulation and nasality in cleft palate speakers, Ph.D. dissertation, Pennsylvania State Univ., 1954.

# Effect of Loudness Recruitment on Delayed Speech Feedback

EARL R. HARFORD

JAMES F. JERGER

There is presently little experimental evidence to show that hearing-impaired subjects behave like normal subjects at identical sensation levels of delayed speech feedback. The results of two investigations (*1*, *2*) suggest that hearing-impaired individuals with recruitment will show greater effects on delayed speech feedback than normal listeners, possibly because of the presence of recruitment. Indeed, theoretical considerations concerning the characteristics of loudness recruitment in pathological ears suggest that individuals manifesting this type of hearing impairment may show reactions to delayed speech feedback at lower sensation levels than either normals or hearing-impaired persons without recruitment. Thus, it may be fallacious to expect that all individuals with hearing impairment will behave like persons with normal hearing under exposure to amplified speech feedback.

Earl R. Harford (Ph.D., Northwestern University, 1958) is Assistant Professor of Otolaryngology, Faculty of Medicine, McGill University. Montreal, and Director, Speech and Hearing Clinic, Royal Victoria Hospital. James F. Jerger (Ph.D., Northwestern University, 1954) is Assistant Professor of Audiology, School of Speech, and Instructor in Otolaryngology, Medical School, Northwestern University. A paper, based on the contents of this article, was presented at the 1958 Convention of the American Speech and Hearing Association, New York.

## *Procedure*

*Experimental Design.* In view of the above considerations an experiment was designed to evaluate the possible contribution of loudness recruitment to the delayed speech feedback effect.

The basic design involved the comparison of delayed speech feedback behavior in three groups: (a) 10 hearing-impaired subjects with recruitment, (b) 10 hearing-impaired subjects without recruitment, and (c) 10 normal listeners.

The selection of subjects for the first group, hearing-impaired with recruitment, presented a problem. In order to afford a fair evaluation of delayed speech feedback, it was deemed necessary to limit all audiometric contours to those that conformed to a pattern of relatively flat and bilaterally symmetrical loss. This selection criterion precluded the use of any direct test of loudness recruitment as a basis for assigning subjects to the first group. It was decided, therefore, to include in this group only subjects whose hearing loss had been medically diagnosed as due to labyrinthine hydrops. Since the literature overwhelmingly indicates that loudness recruitment is a concomitant of this etiology, it was simply assumed that all subjects in this group had loudness recruitment, by virtue of the medi-

cal diagnosis. Selection of subjects for the second group, hearing-impaired without recruitment, seemed to be a less serious problem. This group was composed entirely of subjects with bilateral, medically-diagnosed clinical otosclerosis. The literature clearly indicates that conductive loss of this type is not accompanied by loudness recruitment. For the third group, normal listeners, young adult college students were employed.

Ideally, a comparison of delayed speech feedback behavior in these three groups would be sufficient. In reality, however, the problem is complicated by the fact that hearing-impaired subjects with recruitment, especially those with labyrinthine hydrops, tend to differ from hearing-impaired subjects without recruitment with respect to their ability to understand speech, a variable which could, conceivably, affect performance under delayed speech feedback independent of the presence or absence of recruitment.

In order to assess the effects of these two variables, recruitment and speech discrimination loss, separately, the present design included a fourth group, 10 hearing-impaired subjects with recruitment but without speech discrimination loss. This condition was achieved by creating an artificial hearing loss in normal listeners through the use of binaurally presented broad-band thermal masking noise. Again, the literature clearly indicates the presence of behavior like recruitment accompanying the hearing loss due to a masking noise.

Finally, during the progress of the experiment, it became necessary to add a fifth group, 10 subjects with hearing impairment created by artificially obstructing the external canals. The need for this fifth group became apparent as results obtained on the otosclerotic subjects were analyzed. It was anticipated that, by virtue of the nature of the hearing loss, otosclerotics would behave like normal listeners at comparable feedback sensation levels. It was observed, with some consternation, however, that the otosclerotics showed more marked reaction to feedback than either the labyrinthine hydrops or masked normal groups. Reflection on this unexpected difference between otosclerotics and normals brought two possible explanations. First, the effect might simply be due to the age difference between the two groups. The

TABLE 1. Summary of sex distribution, mean age in years, mean air conduction (AC) and bone conduction (BC) loss in db (500–2000 cps average), mean binaural spondee threshold loss in db, and mean binaural speech discrimination (PB) score, for each of five experimental groups.

| *Group* | *N* | | *Age* | *Loss* | | | *PB Score* |
|---|---|---|---|---|---|---|---|
| | *Male* | *Female* | | *AC* | *BC* | *Spondee* | |
| Normal | 1 | 9 | 19.7 | —* | —* | –1.6 | 96.6% |
| Hydrops | 6 | 4 | 45.9 | 40.4 | 37.8 | 44.0 | 78.8% |
| Masked Normal | 2 | 8 | 20.5 | —* | —* | 42.2 | 95.2% |
| Otosclerotic | 4 | 6 | 45.4 | 38.4 | 9.1 | 41.4 | 96.8% |
| Plugged Normal | 3 | 7 | 42.9 | —* | —* | 27.4 | 96.8%† |

* No pure-tone threshold tests administered. Screened at +10 db.
† Tested for speech discrimination with ear canals unplugged.

mean age for normals was 19.7 years, and the mean age for otosclerotics was 45.4 years. Perhaps, it was thought, susceptibility to the effects of delayed sidetone is much less in young than in middle-aged adults, irrespective of the presence or absence of the conductive hearing loss. The second possibility was that the discrepancy between normals and otosclerotics might simply be due to the presence of a hearing loss, per se, with or without recruitment. It seemed possible that any kind of hearing loss at all, conductive or sensorineural, has the effect of increasing the delayed speech feedback effect. A fifth group of 10 subjects was added to rule out both of these possibilities simultaneously. To this end, middle-aged normal-hearing adults, purposely chosen so that their mean age would roughly approximate that of the otosclerotic group, were tested both with and without wax-impregnated cotton plugs in the external ear canals. Both external canals were plugged in order to produce an artificial hearing loss not accompanied by recruitment. This group was also tested without the external canals plugged in order to isolate the possible age effect from the possible hearing loss effect. In other words, if the behavior observed in the otosclerotics were due simply to their age, the effect would be expected in this fifth group of middle-aged normals whether their external canals were plugged or not. If, on the other hand, the effect observed in the otosclerotics were due simply to the effect of a hearing loss, per se, it would be expected to appear in the fifth group only under the plugged condition.

For each subject in each group, the delayed speech feedback effect was measured at feedback sensation levels of 0, 10, 20, 30, 40, and 50 db. The specific method used to quantify the feedback effect was the flashing-num-

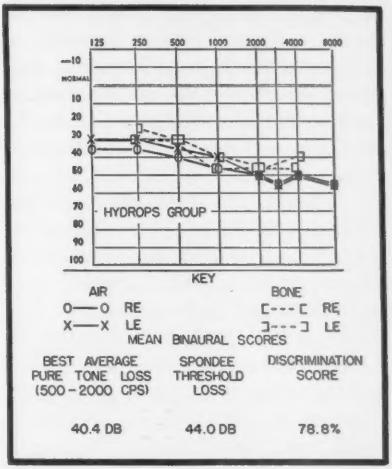



FIGURE 1. Mean pure-tone and speech losses and mean discrimination score for hydrops group.

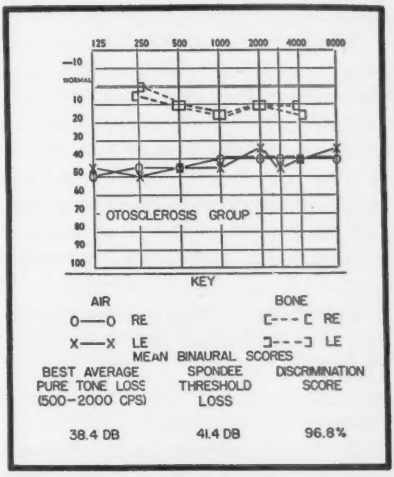



FIGURE 2. Mean pure-tone and speech losses and mean discrimination score for otosclerosis group.

bers technique first described by Butler and Galloway.[1]

*Subjects.* Each experimental group contained 10 subjects, carefully selected on the basis of medical diagnosis[2] or audiological findings or both. For each of the five groups, the sex distribution, mean age, mean binaural spondee threshold losses, and mean binaural discrimination scores were determined, and for the hydrops and otosclerotic groups, the mean best binaural pure-tone average hearing losses for the three speech frequencies (500, 1000, and 2000 cps) by both air and bone conduction (see Table 1). The mean pure-tone threshold hearing levels for the labyrinthine-hydrops and otosclerotic groups are presented in Figures 1 and 2, respectively.

*Apparatus.* A standard pure-tone audiometer was used to obtain the hearing level for each hearing-impaired subject and to screen each subject with normal hearing. A commercial speech audiometer (Grason-Stadler, Model 162) was used for all speech testing. The spondee threshold was established with CID revised recorded spondaic word lists (W-1). Speech discrimination was tested using CID revised recorded phonetically-balanced word lists (W-22). Periodic sound pressure measurements were made over the testing

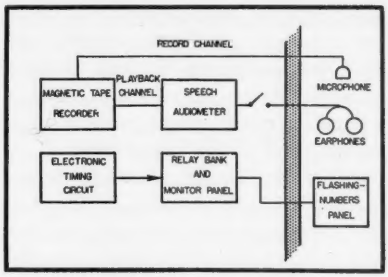



FIGURE 3. Simplified block diagram of the experimental apparatus.

period to ensure the stability of the signal delivered by the earphones.

Figure 3 shows a simplified block diagram of the apparatus used for the delayed speech feedback testing. The path of the circuit originated from a message spoken into a microphone (American, type B-22) fed into a magnetic tape recorder (Magnecord, Model PT6-AH) with fixed record and playback heads. The delay period between the recorded message and the playback message was .167 sec. The output from the tape recorder was fed into the speech audiometer, then through a manual switch, finally to a set of earphones (Telephonic, type TDH-39-10 Z) mounted in MX41/AR cushions. A broad-band thermal noise, incorporated in the speech audiometer, was used for the masked-normal group. The acoustic spectrum of the noise signal was essentially determined by the frequency response characteristic of the earphone. Both speech and noise signals were always delivered binaurally-in-phase. The speech audiometer was used to control the intensity of the delayed speech feedback and, when necessary, to deliver thermal noise to the earphones.

Isolated from the delayed speech feedback apparatus was the flashing-

[1] R. A. Butler and T. F. Galloway in a paper read at the American Speech and Hearing Association convention, Cincinnati, 1957, entitled 'A comparison of the performance of normal hearing and hard of hearing persons on the delayed speech feedback task.'

[2] All subjects in the hydrops and otosclerosis groups were selected from the files of the Department of Otolaryngology, Northwestern University, and from the files of three otologists in private practice in Chicago.

numbers equipment. A multi-vibrator triggered a telephone-type stepping relay at a rate of two steps per second. The relay, in turn, completed a circuit at each of 80 steps, briefly illuminating one of five two-digit numbers on a panel located 4.5 feet in front of the subject's chair and in full view of the subject. These numbers, from the subject's left to right, were 24, 31, 58, 63, and 82, each number being three inches high and centered directly above a six-volt amber light bulb. When a light flashed, the speaker was to say the number above it. The randomization of the stimulus numbers, therefore, was achieved by randomization of the order in which the lights were flashed. On a monitoring panel in full view of the tester, but not seen by the subject, identical numbers were illuminated simultaneously.

Except for the earphones, microphone, and subject's stimulus panel, all testing equipment was located in the control room, isolated from the test chamber.

*Testing.* The over-all testing sequence for each of the 50 subjects in the present investigation consisted of three main procedures: (a) a brief interview and pure-tone testing; (b) speech audiometric testing; and, (c) delayed speech feedback testing.

Each subject in the normal group, masked-normal group, and the plugged-normal group was screened at a hearing level of 10 db, across the frequency range from 125 through 8000 cps, and including 3000 cps, on both ears to ensure normal hearing. The subjects with medically-diagnosed labyrinthine hydrops and otosclerosis were given conventional pure-tone audiometric threshold tests, using an ascending technique.

Following the pure-tone testing, the spondee threshold was established using a similar ascending method. Immediately after the determination of the spondee threshold, phonetically balanced words (W-22 recording series) were presented at a sensation level of 30 db re spondee threshold. Discrimination scores were obtained for subjects in the masked-normal group both in quiet and during exposure to broadband thermal noise.

The third and final test procedure consisted of the flashing-numbers delayed-speech-feedback testing technique at levels of 0, 10, 20, 30, 40, and 50 db above the individual subject's spondee threshold. For each subject in the masked-normal and plugged-normal groups, the spondee threshold obtained during the artificially-produced-hearing-loss condition was used as the base of reference from which sensation level was computed.

For the actual test procedure for each subject, a microphone, mounted on a 'third hand,' was placed at a distance approximately six inches from the speaker's mouth. The manual switch controlling the signal to the earphones was placed in the off position and the subject was presented with a practice series of 40 flashes of lights on his panel to which he responded by saying the numbers as the lights flashed. Following the practice run, the actual delayed speech feedback test began.

During each experimental series, the tester continuously adjusted the record gain on the tape recorder to keep the VU meter needle on the speech audiometer at a predetermined level. This maintained relatively constant input to the subject's earphones in spite of changes in vocal intensity resulting

TABLE 2. Summary of median error scores and ranges of individual error scores for each of five groups at six sensation levels of delayed speech feedback.

| *Group* | | *Sensation Level in db* | | | | | |
|---|---|---|---|---|---|---|---|
| | | *0* | *10* | *20* | *30* | *40* | *50* |
| Normal | Median | 4.5 | 5.0 | 6.5 | 9.0 | 10.5 | 17.0 |
| | Range | 16 | 19 | 21 | 26 | 23 | 23 |
| Hydrops | Median | 12.0 | 14.5 | 17.0 | 22.0 | 28.0 | 24.0 |
| | Range | 26 | 28 | 32 | 35 | 34 | 24 |
| Masked Normal | Median | 5.0 | 9.0 | 14.0 | 17.0 | 21.0 | 25.0 |
| | Range | 15 | 33 | 36 | 32 | 28 | 31 |
| Otosclerotic | Median | 9.0 | 20.0 | 20.5 | 26.0 | 26.5 | 33.0 |
| | Range | 25 | 32 | 42 | 34 | 43 | 28 |
| Plugged Normal | Median | 9.0 | 9.5 | 9.0 | 10.0 | 10.5 | 13.5 |
| | Range | 13 | 14 | 14 | 16 | 18 | 20 |

from the Lombard effect and allied phenomena.

At the conclusion of 40 flashes of lights, the relay automatically ceased. The tester stopped the tape recorder and turned off the earphone switch. A switch on the relay bank was thrown to a second random-order series for the light flashes, and at the end of 15 seconds the same procedure was repeated at the identical feedback sensation level. This provided the subject with a brief rest period between the 40th and 41st number presentations at each sensation level. The order in which the six sensation levels were presented was separately randomized for each subject.

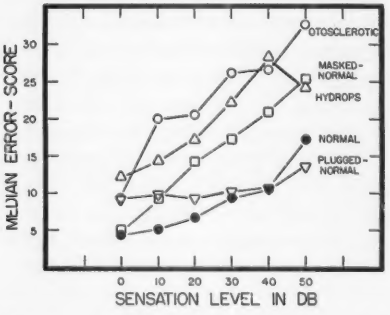



FIGURE 4. Median error-score functions for the five groups.

All subjects in the plugged-normal group were tested twice at each sensation level of delayed speech feedback, once with external ear canals plugged and once with ears unplugged. The order of these test conditions was counter-balanced over the total group.

The total of the numbers repeated correctly by each subject at each sensation level was subtracted from 80, the best possible score. This difference represented the subject's error score at that sensation level. An electric hand counter was used to tabulate correct responses.

## *Results*

Median error scores (Table 2) for the five groups at each sensation level of delayed speech feedback were determined and the results plotted graphically (Figure 4). The curves for the hydrops, otosclerotic, and masked-normal groups are separated from the curves for the normal and plugged-

normal groups by a rather wide gap starting at the 20-db sensation level. The high median error scores obtained for the otosclerotic group were totally unexpected.

### *Discussion*

Viewing these results within the framework of the experimental design suggests that recruitment does indeed accelerate the function relating feedback effect to sensation level. This conclusion seems clear from a consideration of the curves for the hydrops and masked-normal groups. If the factor of lowered speech discrimination scores in the hydrops group were solely responsible for their performance, a comparable effect would not have been expected in the masked-normal group, where there was no impairment in speech discrimination. The fact that both groups behaved in virtually identical fashion suggests the factor common to both groups, loudness recruitment, as the underlying cause.

The unexpectedly high error scores in the otosclerotic group are intriguing, yet difficult to explain, since loudness recruitment is presumably absent in a purely conductive loss. The essentially negative results obtained in the plugged-normal group fail to support either age or the fact of a hearing loss, per se, as possible explanations for the curious results in the otosclerotics.

The clinical implications of the findings reported in this study deserve serious consideration. It has been suggested previously (*3*) that an individual's spondee threshold can be approximated from the lowest intensity level at which speech disturbance can be measured under exposure to delayed speech feedback. The results of the present investigation indicate, however, that this estimate must necessarily differ depending on whether the subject does or does not have a hearing loss accompanied by loudness recruitment. Suppose, for example, that a subject manifests evidence of the delayed speech feedback effect at a hearing level of 70 db (re normal threshold for spondee words). According to the conventional yardstick, his true spondee threshold would be estimated at no more than 30 db. The present results indicate, however, that his loss could be as great as 50 or even 60 db if it were accompanied by recruitment, or even if it were caused by otosclerosis, for that matter.

Tiffany[3] interprets the present findings to mean that, in matters of adjudication, it may be desirable to estimate the actual loss closer to the feedback response level than the behavior of normals would imply. Not to do so would, in his opinion, be to err in favor of the defendant in a damage suit.

### *Summary*

The purpose was to investigate the effect of delayed speech feedback at sensation levels of 0, 10, 20, 30, 40, and 50 db re spondee threshold upon five groups: (a) a normal-hearing group; (b) a labyrinthine-hydrops group (presumably with recruitment); (c) a masked-normal-hearing group (with characteristics analogous to recruitment); (d) an otosclerotic group (without recruitment); and (e) a plugged-normal-hearing group (without recruitment). There were 10 sub-

[3]Personal communication with author, 1959.

jects in each group. High median error scores in the hydrops and masked-normal groups were presumably due to the presence of recruitment. High error scores in the otosclerotic group were apparently due to reasons other than the recruitment phenomenon and remain unexplained.

Some implications of these results for the clinical utilization of delayed speech feedback as a predictor of the true spondee threshold are discussed.

## *References*

1. Butler, R. A., and Galloway, F. T., Factoral analysis of the delayed speech feedback phenomenon. *J. acoust. Soc. Amer.*, 29, 1957, 632-635.

2. Elliott, C. M., Responses of hard of hearing subjects to delayed side-tone. Master's thesis, Vanderbilt Univ., 1955.

3. Hanley, C. N., and Tiffany, W. R., An investigation into the use of electro-mechanically delayed side tone in auditory testing. *J. Speech Hearing Dis.*, 19, 1954, 367-374.

# Language Skills of Profoundly Deaf Adolescent Children

SIDNEY GODA

This study investigated the chief language skills of adolescent pupils at a residential school for the deaf. Four areas were selected for investigation: (a) relationship between chronological age and language skills, (b) relationship between intelligence and language skills, (c) interrelationships among individual language skills, and (d) differences between sexes in development of language skills.

This study was conducted at the (Flint) Michigan School for the Deaf, a state residential school for deaf children. Here training in oral speech is emphasized until the child is 12 years of age; after that, reading and writing receive greater emphasis than oral speech. At the time the children reach senior high school, they are separated into oral and manual groups: the oral group composed of children who have progressed satisfactorily in using oral speech, the manual composed of children who have not progressed satisfactorily in using oral speech.

## *Procedure*

*Subjects.* The subjects selected for this study, all enrolled at the Michigan school, were 56 deaf adolescent children who were without any apparent secondary handicaps. All children selected had losses of 75 db or greater in each frequency of the speech range, 512, 1024, and 2048 cps, and equal or greater losses at the higher frequencies. Each child was either congenitally deaf or deafened by the age of two years. All had entered school at or before the age of six, had received their entire education at a residential school for the deaf, and were placed in oral classes in senior high school.

The 56 subjects, 32 males and 24 females, were divided into age groups as follows: 10, age 12; 13, age 13; 17, age 14; and 16, ages 16 to 18 inclusive.

*Measurement of Skills.* Four language skills were measured: writing, speaking, lipreading, and reading.

Writing. Two samples obtained from each child served as a basis for evaluating written language: (a) a composition on 'Why I Like Summer Better Than Winter' or 'Why I Like Winter Better Than Summer'[1] and (b) 10 sentences in response to 10 individual pictures, one sentence response per picture. The 10 sentences and compositions were evaluated as one unit. The writing measures were collected from groups of six children. Two meetings

Sidney Goda (Ph.D., University of Michigan, 1957) is Director of Speech and Hearing Therapy, Johnstone Training and Research Center, Bordentown, New Jersey. This article is based on a doctoral dissertation completed under the direction of Dr. Warren Ketcham.

[1]Two previous studies (*1*, *8*) report on the adequacy of these topics with deaf children.

were conducted to obtain the samples, the first meeting being devoted to the compositions and the second meeting being devoted to the sentences in reaction to the pictures.

Speaking. Two samples obtained from each subject served as a basis for evaluating spoken language: (a) 10 sentences in response to the same 10 pictures as were used to measure writing ability, one sentence response per picture, and (b) six sentences to tell a story in response to six pictures, one sentence response per picture. There were thus 16 spoken-sentence responses. The samples of spoken language were obtained individually. Because the children were responding orally to the same pictures to which they had formerly responded in writing, an interval of a week was allowed to separate the two presentations in order to foster use of different vocabularies for the spoken and written responses. The child was instructed to repeat each sentence three times while the experimenter wrote down what was said. By being able to see the child and the material to which he was reacting, the experimenter was usually able to understand what the child was saying. In each instance, however, the subject was shown what was written down and asked if it was what he had spoken. The child was given an opportunity to make corrections. When the child could not be understood, he was instructed to write down what he had said. This procedure insured maximum understanding of every word spoken by the child.

Lipreading. A four-part test was constructed to measure lipreading ability: part 1, 10 items, each consisting of two words (an article and a noun); part 2, eight items, each consisting of a three-word phrase; part 3, five items, each consisting of a four-word sentence; and part 4, six items, each consisting of a five- or six-word sentence. There were thus 29 items in all. The test was composed of those words which the majority of children had used at least once in their written compositions. Thus a child was penalized minimally for his inability to spell words or to understand the meanings of the words.

The lipreading test, given to groups of six children, was administered 10 different times. In the administration of the test each item was spoken in a soft voice which the children could not hear because of severe hearing losses. Each item was spoken three times in succession. Then the children were allowed sufficient time to write the item. During the presentation the children were seated around a rectangular table, three on either side, with the natural light from the window at the side of the children and directly over the head of the experimenter who was reading the test material. Six feet was the maximum distance at which any individual child was seated from the speaker. Children with eyeglasses or visual difficulties were placed closest to the speaker.

Reading. Reading ability was measured by the children's scores from the reading portions of the elementary and intermediate batteries of the 1953 revision of the Stanford Achievement Test, administered during the 1953-54 school year.

Intelligence. The IQ of each child was determined by his performance on a nonlanguage intelligence test. All of

the children had been tested between 1952 and 1954 by the school psychologist. Two tests were used, the children between the ages 12-14 being given the Chicago Non-Verbal Test, and the older children being given the Revised Beta Performance Test.

*Scoring Procedures.* Three measures were obtained from two writing samples, two of a quantitative nature and one of a qualitative nature. The quantitative measures, used by numerous experimenters (*6*) in evaluating the language ability of hearing children, were length of composition and average sentence length. The total number of words which the child used yielded the score for the length of the composition. Average sentence length was determined by dividing the total number of words by the number of sentences. The qualitative measure, used for the development of the sentence, was determined by use of an 11-point scale which the writer constructed. The scale ranged from *1*, for a completely unintelligible response showing no concept of a sentence, to *11*, for an easily intelligible, elaborated sentence made up of more than two clauses. Each written sentence was evaluated separately. The average score of all the sentences was the average score for development of written sentence.

The same three measures were used to evaluate spoken language ability: average length of spoken composition, average length of spoken sentence, and average score for development of spoken sentence. The same method which had been used to determine the development of the written sentence was used also to determine the development of the spoken sentence.

In the measurement of lipreading ability, each word correctly understood by the child was graded separately. An article received a score of one-half point, and any other word received a score of one point. The maximum score was 90.

The reading score was the average of the two scores from the reading portions of the Stanford Achievement Test, paragraph meaning and word meaning.

There were thus eight measures: three for written language, three for spoken language, one for lipreading, and one for reading.

TABLE 1. Mean language measures by age groups: 12, 13, 14, and 16 to 18 years, with Ns of 10, 13, 17, and 16, and mean IQ's of 96.5, 96.3, 91.4, and 103.1, respectively.

| *Language Measures* | *Ages* | | | |
|---|---|---|---|---|
| | *12* | *13* | *14* | *16-18* |
| Length of Written Composition in Words | 195.6 | 193.1 | 182.8 | 198.1 |
| Development of Written Sentence (maximum score, 11) | 4.5 | 4.1 | 3.5 | 5.1 |
| Length of Written Sentence in Words | 7.9 | 7.6 | 6.7 | 8.7 |
| Length of Spoken Composition in Words | 101.0 | 99.0 | 83.5 | 108.2 |
| Development of Spoken Sentence (maximum score, 11) | 4.7 | 4.2 | 3.5 | 4.7 |
| Length of Spoken Sentence in Words | 6.0 | 6.2 | 5.5 | 6.8 |
| Lipreading (maximum score, 90) | 37.1 | 37.5 | 28.2 | 53.8 |
| Reading Grade Level | 3.1 | 2.7 | 2.6 | 4.2 |

TABLE 2. Pearson product moment correlation coefficients for all possible pairs of language measures. (With 54 *df*, a correlation of .26 is significant at the 1% level, .34 at the 5% level.)

| *Language Measures* | *2* | *3* | *4* | *5* | *6* | *7* | *8* |
|---|---|---|---|---|---|---|---|
| *1* Length of Written Composition | .57 | .63 | .27 | .51 | .24* | .44 | .49 |
| *2* Development of Written Sentence | | .69 | .59 | .70 | .48 | .59 | .71 |
| *3* Length of Written Sentence | | | .58 | .59 | .56 | .47 | .56 |
| *4* Length of Spoken Composition | | | | .65 | .88 | .42 | .52 |
| *5* Development of Spoken Sentence | | | | | .45 | .43 | .56 |
| *6* Length of Spoken Sentence | | | | | | .40 | .51 |
| *7* Lipreading | | | | | | | .61 |
| *8* Reading Grade Level | | | | | | | |

* Not significant.

## *Results*

Reliabilities of the two measures used for the expressive language skills of writing and speaking were estimated by means of the split-half method employing the Spearman-Brown formula. The reliabilities of the average score for development of written sentence and average length of written sentence were both .90 while the reliabilities of the average score for the development of spoken sentence and average length of spoken sentence were .84 and .92, respectively. The reliability for the lipreading test, calculated in the same manner, was .62. Two possible causes may account for this relatively low reliability: (a) the type of material (words, phrases, and sentences) which made up the lipreading test, since previous research (2) has demonstrated that lipreading scores vary with the type of material, and (b) the changing conditions (particularly lip movements and their visibility) during the 10 administrations of the test.

*Chronological Age and Language Skills.* The means for the eight language measures (Table 1) show a general decrement from the group of 12-year-old children through the next two older groups. However, the oldest group (16 to 18 years) generally scored higher on the language measures.

In order to test the significance of the differences in performance among three of the four age groups (12-year, 13-year, and 14-year), partial correlations were figured between chronological age and the eight language measures with IQ constant.[2] Two negative correlations, average score for development of spoken sentence and total amount of speaking, are significant for these three age groups. For the entire group (all four age levels), however, five correlations between language measures and chronological age are positive and significant. The three measures which did not show positive significant correlations with age for the total group were the ones for the amount of writing, the development of spoken sentence, and the amount of speaking, these latter two being the measures of oral per-

[2]This score was held constant because the correlation coefficient between chronological age and IQ for these three groups was −.30, which is significant at the .10 level.

TABLE 3. Factor loadings of eight language-measure variables on first two factors extracted in centroid factor analysis.

| *Tests* | *Factor A** | *Factor B†* | *Communality* ($h^2$) |
|---|---|---|---|
| Length of Written Composition | .63 | .41 | .5650 |
| Development of Written Sentence | .84 | .10 | .7681 |
| Length of Written Sentence | .80 | .25 | .6500 |
| Length of Spoken Composition | .80 | −.53 | .9209 |
| Development of Spoken Sentence | .77 | .05 | .5954 |
| Length of Spoken Sentence | .73 | −.56 | .8465 |
| Lipreading | .66 | .14 | .4552 |
| Reading Grade Level | .78 | .14 | .6280 |

* General language factor.
† Language confidence factor.

formance which showed significant decreases from the 12-year, to the 13-year, to the 14-year age group.

*Intelligence and Language Skills.* Partial correlations for the three age groups (12-year, 13-year, and 14-year) with chronological age constant revealed five measures significantly related to intelligence: all three measures of written language, average score for development of spoken sentence, and average reading grade level as determined by scores on the reading portions of the Stanford Achievement Test. The reading score is the only measure related significantly to intelligence with the 16-to-18-year age group.

*Interrelationships of the Language Skills.* Pearson product-moment correlation coefficients were computed over the entire age range for all possible pairs of tests in order to estimate the degree of interrelationship of the language skills. The intercorrelations are presented in Table 2. Twenty-seven out of the 28 correlations are positive and significant.

*Underlying Abilities.* A factor analysis was carried out to procure information concerning the underlying abilities accounting for the observed significant correlations among the eight tests. The factoring was done by Thurstone's centroid method to two factors, and the resulting matrix is shown in Table 3. After the removal of these two factors, no further factors were extracted because of the small size of the residuals. Their mean absolute value was .04, and only one was above .10. The communalities listed in Table 3 indicate the amount of variance which the two factors contribute to the tests. With the exception of the lipreading test, these two factors account for over one-half of the variance of each test.

*Sex Differences in Language Skills.* Differences between the sexes[3] were determined for each of the eight language measures. The differences all favored females, but none of these differences was statistically significant as determined by Fisher's *t* test. The measure which showed the largest sex difference was the one for the development of the spoken sentence, significant beyond the 10% level.

[3] No significant differences were found between the sexes in intelligence or chronological age.

*Discussion*

The inconsistent patternings of language skills found with the four age groups in this study (see Table 1) are similar to those found in several other studies. Pintner (7) comments on the erratic pattern of growth in the language of the deaf child. Lyon and others (*4*) found uneven growth in reading, with the largest improvement occurring above grade nine where the average age was 18. The results of two other studies show an uneven patterning in written language. Heider and Heider (*1*) found decreases in sentence length between 13- and 14-year-old children and between 15- and 16-year-old children; Wells (*8*) found seven instances in intermediate grades where the language of older deaf children was inferior to the language of younger deaf children.

The kind of instructional program presented in the particular school used in the present study may account for (a) the significance of decreases in oral language measures from the 12-year to the 13-year to the 14-year age group, (b) the nonsignificance of increases in oral language measures from the youngest to the oldest group, and (c) the significance of increases in the writing and reading measures from the youngest to the oldest group. In the curriculum under discussion, oral language is stressed until the age of 12. After this age less stress is placed on oral language. The child at age 12, then, may have reached a high point in his ability to use oral speech. Increased emphasis on reading and writing after age 12 appears to parallel a greater proficiency in these language skills than in oral language, which is relatively neglected in the educational program of the adolescent children.

Intelligence was found to be more closely related to the language measures for the 12-year, 13-year, and 14-year age groups than for the 16-to-18-year age group. Therefore intelligence would seem to play a more significant role in language development for the younger adolescent deaf child than for the older adolescent deaf child.

The intercorrelations of the language measures as shown in Table 2 indicate a high consistency of interrelatedness. Twenty-seven out of 28 correlations are positive and significant. There is apparently a tendency for the deaf child who is superior in one language skill to be relatively superior in all language skills and for the deaf child who is inferior in one language skill to be relatively inferior in all language skills. Thus the child who excels in reading may be expected, in general, to lipread, write, and speak at a relatively high level while the child who reads at a lower level may be expected, in general, to lipread, write, and speak at a relatively low level.

Further, within the two expressive skills of writing and speaking, there was found to be a significant relationship between the quantity of expression and the quality of expression. The child who uses a relatively greater number of words in his written or oral language may be expected to express himself with relatively longer and more complex sentences. The average length of spoken sentence is, however, not related significantly to the average length of written composition.

The two factors which were extracted in the factor analysis account for over one-half the variance of each

test with the exception of the lipreading test. The Heiders (*2*) have stated that individual factors are responsible for superiority in lipreading. Therefore this language skill may not be affected by any general ability to the same extent as are the other language skills.

Factor A, general language factor, was found in each of the eight tests, suggesting that a single factor underlies all of the language measures. Ketcham (*3*), Williams (*9*), and Yedinack (*10*) suggest that a general language factor underlies the individual language skills for hearing children. Olson (*6*) implies that in language development there is an underlying unity which is expressed in the various structures and functions that are measured.

Factor B, language confidence factor, was found to exist in three of the tests: amount of writing, amount of speaking, and average length of spoken sentence. This factor was present to a lesser extent in the average length of the written sentence. Factor B was positively weighted in the two written tests but negatively weighted in the two oral tests. Negative loading parallels an inhibition of speaking but a positive loading is coincident with a greater quantity of writing.

An analysis of the school environment and test situation may account for the presence of Factor B. The older children generally use sign language, relying very little on oral speech for purposes of communicating their thoughts. Therefore the oral speaking test may have been fear provoking to the subjects since it demanded a form of response which was somewhat foreign to them and one which they did not feel adequate in handling. A further element in creating fear was the presence of the experimenter, who was not only a normal hearing person and stranger but also one who could not communicate with or understand sign language. Since this test was conducted on an individual basis, the child did not receive any outside support from others in his school environment, which would have served to minimize the fear. Being fearful of speaking and lacking confidence in their ability to speak, the children spoke as little as possible, with as short sentences as possible.

The written tests, however, required a response for which the children felt adequate. They had practiced writing in the classrooms and were confident of their ability to express themselves with written language. Further, groups of children were present and responded at the same time. Any fear produced in an individual child by the presence of the experimenter was probably allayed to a great extent by the presence of other children from the same school environment. Because of a feeling of confidence in their ability to express themselves with written language, the children wrote relatively long compositions with relatively long sentences.

The differences in language measures all favor the female, but these differences lack statistical significance. These findings are comparable to results found with normal hearing children as summarized by McCarthy (*5*, pp. 492-630).

## *Summary*

The chief language skills of 56 profoundly deaf adolescent children at a residential school were studied. The 56 subjects, 32 males and 24 females, were divided into four age groups. Four language skills were measured:

writing, speaking, lipreading, and reading.

The following conclusions are drawn: (a) because of the consistent interrelationships of the language skills, the prediction may be made that the deaf child who is superior in one will generally be superior in all of the skills, and the deaf child who is inferior in one will generally be inferior in all of the skills; and (b) the quantitative and qualitative aspects of expressive language appear to be related; thus the child who uses a relatively larger number of words in his speaking, and writing will generally express himself with relatively longer and more complex sentences.

## *References*


1. Heider, F. K., and Heider, Grace M., A comparison of sentence structure of deaf and hearing children. *Psychol. Monogr.*, 52, 1940, 42-103.
2. Heider, F. K., and Heider, Grace M., An experimental investigation of lip-reading. *Psychol. Monogr.*, 52, 1940, 125-153.
3. Ketcham, W. A., Experimental tests of principles of developmental anatomy and neuroanatomy as applied to the pedagogy of reading. Ph.D. thesis, Univ. Michigan, 1951.
4. Lyon, V. W., Stein, S., Levin, A., Johnson, Emily, Fogwell, Dorothy, and Brown, A. W., Report of the 1931 survey of the Illinois school for the deaf. *Amer. Ann. Deaf*, 78, 1933, 157-175.
5. McCarthy, Dorothea, Language development in children. In L. Carmichael (Ed.), *Manual of Child Psychology*. New York: John Wiley, 1954.
6. Olson, W. C., *Child Development*. Boston: D. C. Heath, 1949.
7. Pintner, R., The measurement of language ability and language progress of deaf children. *Volta Rev.*, 20, 1918, 755-766.
8. Wells, C. O., *The Development of Abstract Language Concepts in Normal and in Deaf Children*. Chicago: Univ. Chicago Libraries, 1942.
9. Williams, H. M., An analytical study of language achievement in preschool children. *Univ. Ia Stud. Child Welf.*, 13, 2, 1937, 9-18.
10. Yedinack, Jeanette G., A study of the linguistic functioning of children with articulation and reading disabilities. *J. genet. Psychol.*, 74, 1949, 23-59.

# Language Skills of Male and Female Kindergarten Children

HARRIS WINITZ

Previous language investigations have generally shown girls to be superior to boys on several language skills (*4, 5, 6, 7, 11, 20, 25*). The differences, however, have been small and have rarely been statistically significant. In addition, the differences favoring girls have not been consistent with regard to the age level or language skill investigated. For example, Templin (*20*) reported a statistically nonsignificant difference favoring boys on all language skills investigated at age three (the first age level of her study), but McCarthy (*11*) found girls to be superior to boys on these skills at age three. With regard to articulation skills Davis (*5*) and Templin (*18*) found that girls achieved mastery of articulation earlier than boys, but Roe and Milisen (*14*) reported no statistically significant differences between boys and girls on the number of erroneously uttered sounds at comparable age levels (grades 1 through 6).

A summary of findings regarding sex differences on each of six commonly employed language measures reported in the literature (*2, 4, 6, 7, 11, 14, 15, 17, 20*) is presented in Table 1. These differences favor girls almost two to one, but the differences are statistically significant in only a few instances.

In many of these investigations, however, both intelligence and socioeconomic status, two variables which are known to be related to the linguistic development of children, have not always been equivalent for male and female children. In McCarthy's study (*11*) the girls were more intelligent and were selected from higher socioeconomic levels than the boys. Fisher (7) reported IQ scores for the girls higher than those for the boys. In some instances Templin's (*20*) IQ-equivalents are higher for the girls than for the boys.

Population differences between sexes either in intelligence as measured by IQ scores or in socioeconomic status should presumably not exist because (a) intelligence tests have been designed (*16, 22, 24*) so as not to discriminate between sexes, and (b) it does not seem reasonable to assume that more girls than boys are found in higher socioeconomic classes.

If real differences in language development exist between sexes, it becomes

Harris Winitz (Ph.D. University of Iowa, 1959) is Research Associate in the Bureau of Child Research, University of Kansas. This article is based on a doctoral dissertation completed under the direction of Dr. Frederic L. Darley. This investigation was supported by a scholarship awarded by Kappa Delta Phi under the auspices of the American Speech Correction Fund.

TABLE 1. Summary of findings regarding sex differences on each of six measures reported in the literature (*2, 4, 6, 7, 11, 14, 15, 17, 20*) on language skills of children. Entries under *Age Levels* indicate the total number of tests; the remaining entries show the distribution of results.

| *Measure* | *Age Levels* | *Superior* | | | *Significantly Superior* | |
|---|---|---|---|---|---|---|
| | | *Boys* | *Girls* | *Neither* | *Boys* | *Girls* |
| Mean Length of Response | 39 | 13 | 22 | 4 | 0 | 1 |
| Variability of Response | 8 | 3 | 5 | 0 | 0 | 0 |
| Number of Different Words | 15 | 5 | 10 | 0 | 1 | 0 |
| Structural Complexity Score | 8 | 2 | 6 | 0 | 0 | 1 |
| Articulation | 36 | 14 | 21 | 1 | 0 | 3 |
| Vocabulary | 18 | 7 | 11 | 0 | 0 | 1 |
| Total | 124 | 44 | 75 | 5 | 1 | 6 |

important to understand those environmental, physiological, and mental factors which differentially affect the sexes. A knowledge of such factors may lead to a greater understanding of the processes of language development and retardation. However, before attempts are made to explain sex differences, it appears necessary to establish, using large samples of subjects, whether or not significant sex differences can be demonstrated.

The present investigation was planned in an effort to answer the following question: are the reported relatively small language differences which favor females real differences which would be significant with large samples of subjects or are they chance differences?

## Procedure

*Selection of Subjects.* The subjects were selected from children who were either living in Iowa City homes or residing in towns immediately surrounding Iowa City and who were expected to enter the kindergarten classes of the Iowa City schools in September, 1958. Seventy-five girls and 75 boys were selected randomly from lists of entering children made available through the office of the Iowa City Superintendent of Schools.

*Sampling Procedure.* The proportion of kindergarten children scheduled to enroll at each school, out of the total number of children expected to enter all Iowa City kindergarten classes, was determined; the appropriate proportion of the total sample of 150 was then selected from the kindergarten enrollees of each school. An equal number of boys and girls was selected from each school, exclusive of children of University of Iowa students, rural children, and twins. Children of University of Iowa students living in temporary married-student housing were excluded because it was difficult to classify these families on a socioeconomic scale and it was felt that adequate testing facilities would not be available in their homes. Rural children were eliminated because difficulties encountered in obtaining the cooperation of these families did not warrant their inclusion. Twins were eliminated because a disproportionate number of twins in either sex group might substantially affect scores obtained on certain language measures.

Subjects were restricted to physically normal white children who were from monolingual homes, were considered to be normal in intelligence (that is, with Full Scale IQ's of 70 or above on the Wechsler Intelligence Scale for Children), were currently considered by their parents not to stutter, and were found to have normal hearing. The children's hearing was tested by a portable Microtone ADC audiometer at 4000 cycles at an intensity level of 20 db. This screening procedure, as reported by House (*9*) and House and Glorig (*10*), accomplishes essentially the same purpose as any presently used auditory screening method.

*Nonlanguage Measures.* Information was obtained upon which a comparison of the boys and girls could be based, as indicated below:

Chronological Age: The CA of each child was determined to the nearest month.

Intelligence: Each child was administered the Wechsler Intelligence Scale for Children (WISC) (*24*). From this test three IQ scores were obtained: Verbal Scale IQ, Performance Scale IQ, and Full Scale IQ.

Socioeconomic Status: The socioeconomic class of the child's family was estimated through the use of the Index of Status Characteristics (*23*).

Family Constellation: The family constellation of each child was classified into one of the following five categories: (a) only child, a child having no older or younger sibling within six years of his age; (b) singleton with older siblings, a child having one or more older siblings within six years of his age but having no younger sibling within six years of his age; (c) singleton with younger siblings, a child having one or more younger siblings within six years of his age but having no older sibling within six years of his age; (d) singleton with older and younger siblings, a child having at least one older and one younger sibling within six years of his age.

*Language Measures.* The following language measures were administered to all subjects:

Verbalizations:[1] Fifty verbal responses were elicited in the following manner: the examiner presented one Children's Apperception Test (*3*) card at a time and engaged the child in conversation with neutral comments such as 'Tell me about this picture.' Analyses were made of mean length of response, number of one-word responses, mean number of words in the five longest responses, mean standard deviation, number of different words, and structural complexity. Methods of scoring were similar to those used by McCarthy (*11*), Davis (*4*), and Templin (*20*). The procedure suggested by Templin (*20*, pp. 81-82) for quantifying structural complexity was followed.

Articulation Measures: The Templin Screening Test of Articulation (*19*) was administered. An attempt was made to elicit each sound spontaneously in response to a pictorial stimulus. When this procedure failed, the child was asked to repeat the word after the examiner. These two methods have been found by Templin (*18*) to give essentially the same results. Each item was scored as correct or incorrect.

Vocabulary: The Ammons Full-Range Picture Vocabularly Test, Form A (*1*), was administered.

[1]See Appendix for procedures of elicitation and scoring not specified by previous investigators.

TABLE 2. Means, standard deviations, and ranges of nonlanguage measures for boys, girls, and sexes combined. Values in parentheses are normative means and standard deviations from *Manual for Wechsler Intelligence Scale for Children.*

| *Measure* | *Boys* | | *Girls* | | *Sexes Combined* | | |
|---|---|---|---|---|---|---|---|
| | *Mean* | *S.D.* | *Mean* | *S.D.* | *Mean* | *S.D.* | *Range* |
| CA (months) | 63.76 | 2.65 | 63.24 | 2.83 | 63.50 | 2.65 | 60–72 |
| Socioeconomic Status | 42.07 | 12.57 | 42.92 | 11.53 | 42.49 | 12.08 | 21–76 |
| WISC Verbal Scale IQ | 100.05 | 13.71 | 100.95 | 9.75 | 100.50 | 11.91 | 69–142 |
| | (99.5) | (13.8) | (99.8) | (13.4) | | | |
| WISC Performance Scale IQ | 98.23 | 12.49 | 102.76 | 10.34 | 100.49 | 11.66 | 72–138 |
| | (98.6) | (16.9) | (101.0) | (13.9) | | | |
| WISC Full Scale IQ | 99.07 | 12.85 | 101.95 | 9.22 | 100.51 | 11.27 | 72–133 |
| | (99.0) | (15.4) | (100.4) | (13.2) | | | |

Word Fluency Measures: Four measures of word fluency were obtained: rimes, child names, adult names, and thing names. Procedures for administration of these items were adapted from Gewirtz (*8*). The child was given one minute for each of the following six tasks: (a) name as many words as possible that rime with the 'sounds' -ail, -ake, and -ing; (b) name as many children, adults, and things as possible.

*Order and Place of Test Administration.* The tests were administered to each child in two one-hour sessions, as follows: in the first session the Performance and Verbal Scales of the WISC were administered and the word-fluency measures were obtained. In the second session the articulation and vocabulary tests were administered and 50 verbal responses were elicited and recorded. The two testing sessions were completed within three days. Only the examiner, who administered all tests, was present with each subject. Except in the case of two boys, all tests were conducted in the child's home, usually in his bedroom or in a playroom.

*Examiner Reliability.* The reliability of the examiner was determined in the following manner:

Observer Reliability. The Templin Screening Test was administered to six kindergarten children, their responses being recorded by the author and one other trained examiner. Percentages of observer agreement were calculated for each sound for each child. The average percentage of agreement was 93%.

In order to determine how well two observers agree, using the McCarthy-Davis procedure, in identifying a spoken response, the author and one other

TABLE 3. Summary of $t$ tests for evaluation of differences between boys and girls on several nonlanguage measures ($df$ = 148).

| *Measure* | *Mean Difference* | *t* | *Sign. Level* |
|---|---|---|---|
| CA | .52 | 1.14 | .30 |
| Socioeconomic Status | .85 | .43 | .70 |
| WISC Verbal Scale IQ | .90 | .55 | .70 |
| WISC Performance Scale IQ | 4.53 | 2.41 | .02 |
| WISC Full Scale IQ | 2.89 | 1.57 | .20 |

TABLE 4. Chi square testing independence of frequencies of family constellation types for boys and girls.

| *Family Constellation Type* | *Boys* | *Girls* | *Total* |
|---|---|---|---|
| Only Children* | 5 | 6 | |
| Only Older Siblings | 28 | 20 | 48 |
| Only Younger Siblings | 16 | 22 | 38 |
| Older and Younger Siblings | 26 | 27 | 53 |
| Total | 70 | 69 | 139 |

Chi square ($df = 2$) = 2.29; $P = .50$

* Frequencies of only children were excluded in the chi-square computation.

observer recorded on paper the first and last word of 15 remarks from the same six subjects. Percentage of agreement was determined by comparing the first and last words of the 15 responses recorded by the two observers. The average percentage of agreement was 94%. In the investigation proper the author identified each response in the same way and in addition each subject's responses were tape recorded so that a typed script of the complete set of remarks could be prepared.

Scorer Reliability. The self-agreement of the examiner was found for length of response and structural complexity score. From the typed protocols a random selection of the records of 10 subjects (a total of 500 responses) was made and each response was rescored. The obtained percentages of agreement were 99% and 96% for the rescoring of length of response and structural complexity score, respectively. The reliability coefficients in general indicate high examiner agreement.

## *Results*

*Nonlanguage Measures.* The means, standard deviations, and ranges for chronological age, socioeconomic status, and IQ are presented in Table 2. The results of the $t$ tests of significance of sex differences are given in Table 3. The only nonlanguage difference that

TABLE 5. Means, standard deviations, and ranges for verbalization measures, Templin Screening Test of Articulation, Ammons Full-Range Picture Vocabulary Test, and word-fluency measures for boys, girls, and sexes combined.

| *Measure* | *Boys* | | *Girls* | | *Sexes Combined* | | |
|---|---|---|---|---|---|---|---|
| | *Mean* | *S.D.* | *Mean* | *S.D.* | *Mean* | *S.D.* | *Range* |
| Length of Response | 5.16 | 1.52 | 5.61 | 1.76 | 5.39 | 1.66 | 1.7–13.8 |
| Number of One-Word Responses | 8.33 | 5.39 | 7.40 | 3.87 | 7.87 | 4.80 | 1–34 |
| Mean of Five Longest Responses | 13.44 | 5.69 | 16.40 | 7.85 | 14.92 | 7.01 | 4.4–52.8 |
| Mean S.D. | 3.60 | 1.29 | 4.11 | 1.58 | 3.85 | 1.46 | 1.2–16.2 |
| Number of Different Words | 102.24 | 25.44 | 110.29 | 24.98 | 106.27 | 25.53 | 43–181 |
| Structural Complexity Score | 38.43 | 14.56 | 42.37 | 13.96 | 40.40 | 14.39 | 1–73 |
| Templin Test | 40.47 | 10.15 | 41.69 | 9.33 | 41.08 | 9.75 | 7–50 |
| Ammons Test | 25.19 | 5.39 | 25.04 | 4.24 | 25.11 | 4.90 | 14–38 |
| Rimes | 2.69 | 3.16 | 2.41 | 3.00 | 2.55 | 3.16 | 0–12 |
| Child Names | 7.75 | 3.00 | 9.48 | 4.00 | 8.61 | 3.61 | 2–21 |
| Adult Names | 3.36 | 3.00 | 4.23 | 3.16 | 3.79 | 3.16 | 0–14 |
| Thing Names | 12.04 | 5.83 | 11.93 | 5.00 | 11.99 | 5.39 | 0–26 |

TABLE 6. Summary of *t* tests for evaluating differences between boys and girls on the several language measures ($df$ = 148).

| *Measure* | *Mean Difference* | *t* | *Sign. Level* |
|---|---|---|---|
| Length of Response | .45 | 1.64 | .10 |
| Number of One-Word Responses | .93 | 1.21 | .30 |
| Mean of Five Longest Responses | 2.96 | 2.63 | .01 |
| Mean S.D. | .51 | 2.17 | .05 |
| Number of Different Words | 8.05 | 1.94 | .10 |
| Structural Complexity Score | 3.94 | 1.68 | .10 |
| Templin Test | 1.22 | .77 | .50 |
| Ammons Test | .14 | .02 | .90 |
| Rimes | .28 | .55 | .70 |
| Child Names | 1.73 | 2.98 | .01 |
| Adult Names | .87 | 1.71 | .10 |
| Thing Names | .11 | .12 | .90 |

reaches significance at the .05 level is on the WISC Performance Scale IQ. However, the WISC Full Scale IQ scores for the two groups were nearly equivalent and the groups may be considered essentially the same in intellectual ability.

The frequencies of family constellation types for boys and girls are summarized in Table 4. The chi square of 2.29 was not significant, indicating that the distribution of children in each of the family constellation categories employed in this test was independent of sex.

It appears, then, that the groups of boys and girls in this study were essentially equivalent with regard to the nonlanguage variables shown in previous language investigations to be importantly related to the linguistic achievement of children.

*Language Measures.* The means, standard deviations, and ranges of the language measures are presented in Table 5. The results of the *t* tests are given in Table 6. The .05 level of confidence was established as the minimum probability level required for detecting statistical significance.

Verbalizations. As shown in Table 6, the direction of the differences for all six verbalization measures is in favor of the girls. However, the differences are significant in the case of only two measures, mean of the five longest responses and mean standard deviation.

Templin Screening Test of Articulation. The difference in articulation scores favors the girls by a little more than one sound. This difference, however, is not significant.

Ammons Vocabulary Test. The difference in vocabulary scores favors the boys, but the difference is small and nonsignificant.

Word-Fluency Measures. For two measures, Rimes and Thing Names, the differences favor the males and for two measures, Child Names and Adult Names, the differences favor the females. Only the difference for Child Names is significant.

## *Discussion*

*Verbalization Measures.* A more advanced development of girls over boys at five years of age on the verbalization scales is not demonstrated in this study. The obtained values indicate significant differences on two of three variability measures but do not indicate significant differences on the verbalization measures which are generally regarded as of major importance: length of response, structural complexity score, and number of different words. These statistically nonsignificant results indicate considerable overlap between the two

groups on these measures. It would appear that on the basis of the statistical tests employed, the hypothesis of no difference between sexes on the major verbalization measures is tenable. Furthermore, the values presented in Table 5 indicate that the differences between the sexes are not of sufficient magnitude to justify regarding the groups as essentially different in verbalization skills, thus contraindicating (a) investigations attempting to explain sex differences and (b) the clinical practice of considering the verbalization skills of the two groups as different (as suggested by the use of separate norms, for example).

It may be argued, however, that since the *t* values in all instances approached or reached significance, there may be slight differences favoring the girls on all verbalization measures but that these differences would be detectable only with a sample of subjects considerably larger than that employed in this investigation. Both the cost of conducting such an expanded investigation and the possible limited usefulness of the results make the practicability of such an investigation questionable.

The possibility may be suggested that one of the reasons for not obtaining statistically significant differences in favor of the girls was that the examiner was a male. However, three of the 10 persons who administered tests in the Templin investigation were male, and one of the men did a substantial amount of the testing.[2] Since children in the present culture typically interact with members of both sexes, it would seem that the sex of the examiner is not an important source of error.

[2]Personal communication from Dr. Templin, 1959.

In this study the only values which compare closely with those of Templin (*20*) pertain to length of response. Noticeable discrepancies appear between the mean values of the other verbalization measures in this study and the corresponding values in Templin's study. These differences arose in all probability from the use of different stimulus materials and from procedural differences, the most important of which was probably the use of a tape recorder. In previous studies the accuracy of transcribing responses rested entirely with the listener, who had to make rapid judgments concerning fleeting verbalizations. This investigator found in typing the tape-recorded responses that it was usually necessary to replay the responses several times in order to be absolutely certain of what the child had said. In many instances the examiner felt that he had heard a verb, article, or conjunction, but upon replaying he often found that the child had omitted these words. This reason can account for the fact that the structural complexity scores are lower than those reported by Templin.

*Templin Test of Articulation.* On the basis of the findings, the hypothesis of no difference between sexes on the Templin Test appears tenable. Differences favoring females have not been observed as frequently in articulatory skills as in verbalization measures. Roe and Milisen (*14*) and Sayler (*15*), who studied approximately 2000 children from grades one through 12, found no statistically significant differences between boys and girls with respect to the correct production of consonant sounds.

It may be hypothesized that there are no real differences between boys and

girls of five years of age and older with respect to articulatory skills and that the differences previously emphasized may be an overflow generalization from the verbalization measures. However, the results of this study and the studies of Roe and Milisen and Sayler do not permit a statement regarding the comparative articulatory skills of boys and girls below five years of age.

*Ammons Vocabulary Test.* The difference between sexes in vocabulary skills observed in this study, as measured by the Ammons Test, was negligible and statistically nonsignificant. Girls have been generally regarded (*12*) as being more proficient in vocabulary skills than boys. This opinion may have been fostered by the early Smith (*17*) investigation. However, most often the makers of vocabulary tests have attempted to create tests that are equivalent for both sexes. Therefore, it is difficult to say, on the basis of standardized vocabulary tests, whether or not a sex difference exists. If one were to investigate further sex differences with regard to vocabulary acquisition, a logical starting point might be the readministration of the Smith Test to a large sample of boys and girls.

The scores for both the girls and the boys in this study were found to exceed by approximately four words the scores reported by Ammons and Holmes (*2*). As the examiner checked all responses that appeared to be guesses, the difference between the scores in this sample and those in the Ammons and Holmes standardization sample may reflect a true advancement in children's vocabulary skills since the standardization of the test and its administration by Templin approximately 10 years ago. This possibility is, indeed, not surprising. The Revised Stanford-Binet, Form L, is now reported to be yielding higher scores than were reported in the standardization norms (*13*).

*Word-Fluency Measures.* On the basis of the results obtained in this study it seems reasonable to conclude that girls do not possess a general superiority in word fluency. They do appear to be distinctly superior in naming children. Female superiority in naming children may be the result of the tendency for girls to be more sociable than boys (*22*). However, this superiority was not demonstrated in the naming of adults.

## *Summary*

This investigation sought to discover whether previously reported and relatively small differences favoring female children are chance differences or real differences which would be significant with large samples. Subjects were 150 randomly selected, normal, five-year-old children (75 boys, 75 girls) from a single community. The two sex groups were essentially the same with regard to chronological age, IQ, socioeconomic status, and family constellation. On language variables, statistically significant differences favoring the girls were found on two of six verbalization measures and on one of four word-fluency measures. No significant differences were found on the Templin Articulation Test and the Ammons Vocabulary Test. It was concluded that the hypothesis of no language difference between the sexes is tenable in the population of five-year-old children with regard to major verbalization measures, articulatory skills, vocabulary skills, and three of four word-fluency measures.

## *References*


1. AMMONS, R. B., and AMMONS, HELEN S., *The Full-range Picture Vocabulary Test*. Louisville, Ky.: Psychological Test Specialists, 1948.
2. AMMONS, R. B., and HOLMES, J. C., The full-range picture vocabulary test: III. Results for a preschool-age population. *Child Develpm.*, 20, 1949, 5-14.
3. BELLAK, L., and BELLAK, SONYA S., *Children's Apperception Test* (manual, rev. ed.). New York: C.P.S. Co., 1950.
4. DAVIS, EDITH A., The development of linguistic skills in twins, singletons with siblings, and only children from age five to ten years. *Inst. Child Welf., Monogr. Ser.*, No. 14. Minneapolis: Univ. Minn. Press, 1937.
5. DAVIS, IRENE P., The speech aspects of reading readiness. In Chap. 2 in *Newer Practices in Reading in the Elementary School* (17th yearb., dept. elem. school prin., NEA), vol. 17, no. 7, 1938.
6. DAY, ELLA J., The development of language in twins: I. A comparison of twins and single children. *Child Develpm.*, 3, 1932, 179-199.
7. FISHER, MARY S., Language patterns of preschool children. *Child Develpm. Monogr.* (Columbia Univ.) No. 15, 1934.
8. GEWIRTZ, J. L., Studies in word-fluency: I. Its relation to vocabulary and mental age in young children. *J. genet. Psychol.*, 72, 1948, 165-176.
9. HOUSE, H. P., A new instrument and concept for the rapid detection of a possible hearing impairment. *Trans. Amer. Acad. Opthal. Oto-laryng.*, 61, 1957, 228-230.
10. HOUSE, H. P., and GLORIG, A., A new concept of auditory screening. *Laryngoscope*, 67, 1957, 661-668.
11. MCCARTHY, DOROTHEA A., The language development of the preschool child. *Inst. Child Welf., Monogr. Ser.*, No. 4. Minneapolis: Univ. Minn. Press, 1930.
12. MCCARTHY, DOROTHEA A., Language development in children. Chap. 9 in L. Carmichael (Ed.) *Manual of Child Psychology* (2nd ed.). New York: John Wiley, 1954.
13. PASTOVIC, J., and GUTHRIE, G., Some evidence on the validity of the WISC. *J. cons. Psychol.*, 15, 1951, 385-386.
14. ROE, VIVIAN, and MILISEN, R., The effect of maturation upon defective articulation in elementary grades. *J. Speech Dis.*, 7, 1942, 37-45.
15. SAYLER, HELEN K., The effect of maturation upon defective articulation in grades seven through twelve. *J. Speech Hearing Dis.*, 14, 1949, 202-207.
16. SEASHORE, H., WESMAN, A., and DOPPELT, J., The standardization of the Wechsler Intelligence Scale for Children. *J. cons. Psychol.*, 14, 1950, 99-110.
17. SMITH, MADORAH E., An investigation of the development of the sentence and the extent of vocabulary in young children. *Univ. Ia Stud. Child Welf.*, 3, No. 5, 1926.
18. TEMPLIN, MILDRED C., Spontaneous versus imitated verbalization in testing articulation in preschool children. *J. Speech Hearing Dis.*, 12, 1947, 293-300.
19. TEMPLIN, MILDRED C., Norms on a screening test of articulation for ages three through eight. *J. Speech Hearing Dis.*, 18, 1953, 323-331.
20. TEMPLIN, MILDRED C., Certain language skills in children, their development and interrelationships. *Inst. Child Welf., Monogr. Ser.* No. 26. Minneapolis: Univ. Minn. Press, 1957.
21. TERMAN, L. M., and MERRILL, MAUD A., *Measuring Intelligence*. Boston: Houghton Mifflin, 1937.
22. TERMAN, L. M., and TYLER, LEONA E., Psychological sex differences. Chap. 16 in L. Carmichael (Ed.), *Manual of Child Psychology* (2nd ed.). New York: John Wiley, 1954.
23. WARNER, W. L., MEEKER, MARCHIA, and EELLS, K., *Social Class in America*. Chicago: Science Research Assoc., 1949.
24. WECHSLER, D., *Manual for Wechsler Intelligence Scale for Children*. New York: Psychological Corp., 1949.
25. WELLMAN, BETH L., CASE, IDA M., MENGERT, IDA G., and BRADBURY, DOROTHY E., Speech sounds of young children. *Univ. Ia Stud. Child Welf.*, Vol. 5, No. 2, 1931.


## *Appendix*

Procedures of elicitation and scoring not included by McCarthy (*11*), Davis (*4*), or Templin (*20*) are indicated below:

*Recording the Remarks*: The following definition was added: (1) Remarks connected by interjections and conjunctions, such as 'and,' 'um,' 'er,' 'uh,' etc., are considered as separate remarks if the remarks appear to be clearly enumerative. Examples: 'The girl . . . er.' 'The boy . . . um.' 'And the monkey.' Each example was considered a single remark.

*Typing the Remarks:* The following procedures were followed in typing the remarks from the tape recordings:

A. Repetitions. Since the repetition of words and phrases was not considered indicative of linguistic maturity and since it would substantially add to the length of the uttered response, it was decided to exclude repetitions according to the following rules:

1. Word repetitions. Repetition of the same word in a sentence was typed only once: 'He, *he,* he went home.' 'The boy ran *ran* he ran away.' The italicized words were excluded.

2. Phrase repetition. Repetition of the same phrase in a sentence was typed only once: 'He was, *he was* hurt.' 'And he hit, *and he hit* him.' The italicized words were excluded. However, the insertion of at least one different word nullified a phrase repetition: 'And he hit, and he spanked him.' 'They were going, they were having fun.' Word repetition within a phrase repetition did not nullify a phrase repetition: 'And he's, *he's he's, and he's* home.' The italicized words were excluded. Contractions were separated for purposes of this analysis and did not nullify a repetition: 'He's eating, *he is eating* the mouse.' The italicized words were excluded. This procedure was also followed for word repetition.

3. Repetitions not excluded. When a child repeated a word for enumerative purposes and for starting a new 'thought unit,' he was not penalized for repetition: 'That's a bear there, that's a bear there, that's a bear there.' (Child was pointing as he spoke.) 'I think *that that* is a bear.' 'That's what they are, *they're* bears.' Both 'that' and 'they're' were used as part of a new subordinate or independent clause. Where it could not be determined whether a repetition was used for emphasis or constituted a nonfluency, the repetition was excluded.

B. Incompleted words. Words which were not completed by the child were recorded as though they were completed. Examples: 'I th— (think) I know he's going home.' 'He might ma—(make) build a fire.' It was especially important to follow this procedure with regard to the repetition analysis. In many instances the child repeated phrases but did not complete words within the phrase. Example: 'I th—(think) he, I think he went home.' This sentence could hardly be interpreted, as 'I he, I think he went home.'

C. Noises. Noises were typed only when they were considered an integral part of the sentence. 'Ahhh *grrr,* the lion is eating the monkey.' 'The lion is going to *grrr* the monkey.' In the first instance the noise was considered unessential and was excluded as not indicative of the child's linguistic maturity. In the second instance the noise clearly took the part of a verb and was included.

D. Interjections. Interjections not considered dictionary items and functioning solely to connect words or phrases were not included: 'er,' 'um,' etc. However, interjections which served as words were included: 'uhhuh,' 'hmmm.' The first interjection served as a substitute for the affirmative word 'yes' and the second as the negative word 'no.'

E. Colloquialisms and neologisms. All colloquialisms and neologisms were included: 'wham,' 'whoops,' 'yike,' 'ya,' 'yipee,' 'teensy-weensy,' 'naw,' 'yeah,' etc.

F. Incomprehensible remarks. Some remarks were incomprehensible because the child either turned from the microphone, whispered, mumbled, or uttered an inarticulate word. Remarks in which all words were not clearly comprehensible to the examiner were excluded.

*Analyses Made:*

1. Number of different words.

a. Contractions except for 'ain't' were divided into two words. For example, 'didn't' and 'they're' were changed into 'did not' and 'they are.' Both examples were tabulated as two different words.

b. Plurals and possessives were counted as separate words.

c. The variant colloquial forms of 'mamma' were grouped together: 'mom,' 'mommy,' 'ma.' However, the word 'mother' was listed separately. The variant colloquial forms of 'papa' were grouped together: 'pa,' 'pop.' However, the words 'daddy' and 'dad' were listed as separate words and were not included with the colloquial forms of 'papa.' The word 'em' was listed together with the word 'them.'

2. Structural analysis.

a. Sentences beginning with the words 'but,' 'and,' and 'then' were considered complete. Sentences beginning with the conjunctions 'cause,' 'for,' and 'because' were considered incomplete.

b. Compound sentences with one phrase were classified as elaborated and scored with a weight of 4.

# Relationships Between Language and Nonlanguage Measures of Kindergarten Children

HARRIS WINITZ

In a previous investigation (*6*) boys and girls of kindergarten age were compared on several language and nonlanguage measures. The present report examines the relationships among the major language and nonlanguage variables for the children who served as subjects in that investigation.

## *Procedure*

A complete account of the subjects, measures, and procedures used in the collection and analysis of the data has been given by Winitz (*6*). In the present investigation Pearson product-moment correlations were computed between the major language and nonlanguage variables for boys, girls, and boys and girls combined. The Fisher *z* test was employed to determine the statistical significance of differences between correlations computed separately for boys and girls between any two variables. The nonlanguage variables were socioeconomic status and three IQ's as measured by the Wechsler Intelligence Scale for Children (WISC): Verbal Scale IQ, Performance Scale IQ, Full Scale IQ. The language variables were verbalization measures (mean length of response, number of one-word responses, mean of the five longest responses, number of different words, structural complexity score, and mean standard deviation), Templin Screening Test of Articulation score, size of vocabulary as measured by the Ammons Full-Range Picture Vocabulary Test, and four fluency measures (rimes, child names, adult names, and thing names).

## *Results and Discussion*

Pearson product-moment correlations among the major language and nonlanguage variables are presented in Tables 1, 2, and 3. A brief discussion of some of the correlations follows.

*Socioeconomic Status.* The correlations between the scores for socioeconomic status, based upon the Index of Status Characteristics (*4*), and 14 language and nonlanguage variables are small and for the most part do not differ significantly from zero (Table 1). The only correlations significant at the .05 level (using data for boys and girls combined) are between socioeconomic status and the WISC Verbal Scale IQ, Performance Scale IQ, Full Scale IQ,

Harris Winitz (Ph.D., University of Iowa, 1959) is Research Associate in the Bureau of Child Research, University of Kansas. This article is based on a doctoral dissertation completed under the direction of Dr. Frederic L. Darley. This investigation was supported by a scholarship awarded by Kappa Delta Phi, under the auspices of the American Speech Correction Fund.

TABLE 1. Pearson product-moment correlations computed between the major language and nonlanguage measures for boys (B), girls (G), and boys and girls combined (C).

| | *Socio-economic Status* | *WISC IQ* | | |
|---|---|---|---|---|
| | | *Verbal* | *Perf.* | *Full Scale* |
| *WISC Verbal IQ* | | | | |
| (B) | .26* | | | |
| (G) | .27* | | | |
| (C) | .26† | | | |
| *WISC Performance IQ* | | | | |
| (B) | .38 | .57 | | |
| (G) | .13 | .42 | | |
| (C) | .26 | .51 | | |
| *WISC Full Scale IQ* | | | | |
| (B) | .35 | .90 | .86 | |
| (G) | .24 | .85 | .83 | |
| (C) | .30 | .88 | .85 | |
| *Length of Response* | | | | |
| (B) | .18 | .20 | .16 | .21 |
| (G) | .13 | .09 | .13 | .12 |
| (C) | .15 | .15 | .16 | .18 |
| *Number of One-Word Responses* | | | | |
| (B) | .00 | −.42 | −.21 | −.37 |
| (G) | −.19 | −.19 | −.20 | −.27 |
| (C) | −.07 | −.35 | −.22 | −.33 |
| *Mean of Five Longest Responses* | | | | |
| (B) | .18 | .23 | .13 | .21 |
| (G) | .11 | .03 | .09 | −.04 |
| (C) | .12 | .13 | .05 | .11 |
| *Number of Different Words* | | | | |
| (B) | .09 | .29 | .06 | .22 |
| (G) | .23 | .09 | .05 | .08 |
| (C) | .15 | .21 | .09 | .18 |
| *Structural Complexity Score* | | | | |
| (B) | .17 | .26 | .13 | .23 |
| (G) | .01 | .17 | .20 | .22 |
| (C) | .09 | .23 | .18 | .24 |
| *Templin Screening Test of Articulation* | | | | |
| (B) | .09 | .32 | .29 | .34 |
| (G) | .20 | .34 | .22 | .33 |
| (C) | .12 | .33 | .27 | .34 |
| *Ammons Full-Range Picture Vocabulary Test* | | | | |
| (B) | .18 | .76** | .49 | .72** |
| (G) | .21 | .35 | .52 | .53 |
| (C) | .19 | .61 | .49 | .64 |
| *Rimes* | | | | |
| (B) | .25 | .37** | .39 | .43** |
| (G) | .14 | .02 | .25 | .10 |
| (C) | .20 | .23 | .29 | .29 |

(Continued next column)

*Throughout table, for boys (B) and girls (G) separately, $r_{.05}$ (nearest tabled $df$=70) = .23, $r_{.01}$=.30.

†Throughout table, for boys and girls combined (C), $r_{.05}$ (nearest tabled $df$=150) =.16, $r_{.01}$=.21.

**Difference between $r$s, calculated for boys and girls separately, significant at or beyond .05 level as determined by Fisher $z$ Test.

Table 1 (Continued; for footnotes, see preceding column)

| | *Socio-economic Status* | *WISC IQ* | | |
|---|---|---|---|---|
| | | *Verbal* | *Perf.* | *Full Scale* |
| *Child Names* | | | | |
| (B) | .08 | .21 | .13 | .21 |
| (G) | .10 | .29 | −.06 | .14 |
| (C) | .08 | .24 | .07 | .19 |
| *Adult Names* | | | | |
| (B) | −.05 | .27 | .16 | .24 |
| (G) | .10 | .07 | .12 | .10 |
| (C) | .02 | .18 | .16 | .19 |
| *Thing Names* | | | | |
| (B) | .09 | .44 | .42 | .49 |
| (G) | .14 | .37 | .24 | .36 |
| (C) | .12 | .41 | .34 | .43 |

vocabulary size, and rimes; none of the correlations exceeds .30.

Correlations of this small magnitude are probably the result of the homogeneous social status of Iowa City families. The score derived from the use of the Index of Status Characteristics is based upon ratings of occupation type and level, source of income, house size and type, and area lived in. The average Iowa City family lives in a small or medium-sized house in an average residential area. A large proportion of the fathers are salaried. Perhaps the only item that might importantly discriminate among families in Iowa City would be the occupation of the father. If correlations were made between occupation of father and the various language and nonlanguage variables, a higher relationship might be shown.

*Intelligence.* The intercorrelations of the WISC IQ's (Table 1) compare favorably with the intercorrelations reported by Wechsler (*5*) in the standardization sample for children of 7.5 years. (He does not report intercorrelations for the 5.5-year group.) In the present study the correlations, using data for boys and girls combined, are .51 between Performance Scale IQ and

Verbal Scale IQ, .88 between Verbal Scale IQ and Full Scale IQ, and .85 between Performance Scale IQ and Full Scale IQ; the corresponding values reported by Wechsler are .60, .90, and .89, respectively. In the present investigation differences between intercorrelations of the WISC IQ's calculated separtely for boys and girls were not statistically significant.

*Verbalizations.* The correlations (Table 1) between the verbalization scores and various nonlanguage measures are

TABLE 2. Pearson product-moment correlations computed between the major language measures for boys (B), girls (G), and boys and girls combined (C).

| | *Measures* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* |
| *1. Length of Response* | | | | | | | | |
| (B) | .69*† | .55 | .22 | .23 | .00 | .14 | .07 | .13 |
| (G) | .44* | .44 | −.03 | .10 | .17 | .10 | .14 | .15 |
| (C) | .56** | .50 | .10 | .16 | .10 | .14 | .12 | .15 |
| *2. Number of Different Words* | | | | | | | | |
| (B) | | .53 | .21 | .30 | .17 | −.01 | .28 | .15 |
| (G) | | .59 | .06 | .03 | .03 | −.03 | .06 | .16 |
| (C) | | .57 | .14 | .18 | .10 | .02 | .18 | .15 |
| *3. Structural Complexity Score* | | | | | | | | |
| (B) | | | .28 | .28 | .08 | .17 | .11 | .25 |
| (G) | | | .29 | .30 | .01 | −.10 | .12 | .05 |
| (C) | | | .29 | .28 | .04 | .06 | .13 | .16 |
| *4. Templin Screening Test of Articulation* | | | | | | | | |
| (B) | | | | .33 | .18 | .23 | .29 | .41 |
| (G) | | | | .24 | .15 | .16 | .08 | .24 |
| (C) | | | | .28 | .16 | .20 | .20 | .33 |
| *5. Ammons Full-Range Picture Vocabulary Test* | | | | | | | | |
| (B) | | | | | .36 | .18 | .12 | .45† |
| (G) | | | | | .15 | .01 | .13 | .14 |
| (C) | | | | | .27 | .08 | .12 | .32 |
| *6. Rimes* | | | | | | | | |
| (B) | | | | | | .11 | .07 | .25 |
| (G) | | | | | | .00 | .25 | .15 |
| (C) | | | | | | .04 | .15 | .21 |
| *7. Child Names* | | | | | | | | |
| (B) | | | | | | | .16 | .39 |
| (G) | | | | | | | .34 | .38 |
| (C) | | | | | | | .29 | .36 |
| *8. Adult Names* | | | | | | | | |
| (B) | | | | | | | | .28 |
| (G) | | | | | | | | .06 |
| (C) | | | | | | | | .17 |
| *9. Thing Names* | | | | | | | | |

* Throughout table, for boys (B) and girls (G) separately, $r_{.05}$ (nearest tabled $df$=70) = .23, $r_{.01}$=.30.
† Difference between $r$s, calculated for boys and girls separately, significant at or beyond .05 level as determined by Fisher $z$ test.
** Throughout table, for boys and girls combined (C), $r_{.05}$ (nearest tabled $df$=150) = .16, $r_{.01}$= .21.

TABLE 3. Pearson product-moment correlations computed between four verbalization measures for boys (B), girls (G), and boys and girls combined (C).

| Measure | *Number of One-Word Responses* | *Mean of Five Longest Responses* | *Mean S. D.* |
|---|---|---|---|
| *Length of Response* | | | |
| (B) | −.62* | .54† | .87† |
| (G) | −.41* | .21 | .54 |
| (C) | −.52** | .35 | .68 |
| *Number of One-Word Responses* | | | |
| (B) | | −.51† | −.38 |
| (G) | | −.13 | −.28 |
| (C) | | −.31 | −.33 |
| *Mean of Five Longest Responses* | | | |
| (B) | | | .64† |
| (G) | | | .31 |
| (C) | | | .46 |

* Throughout table, for boys (B) and girls (G) separately, $r_{.05}$ (nearest tabled $df$=70) = .23, $r_{.01}$= .30.

† Difference between $r$s, calculated for boys and girls separately, significant at or beyond .05 level as determined by Fisher $z$ test.

** Throughout table, for boys and girls combined (C), $r_{.05}$ (nearest tabled $df$=150) = .16, $r_{.01}$ = .21.

not high. For example, the sexes-combined correlations between length of response and socioeconomic status and WISC Full Scale IQ are .15 and .18, respectively. The correlations (Table 2) between the verbalization scores and the other language measures also are low. The highest sexes-combined coefficient obtained was .29 (between structural complexity score and Templin Screening Test of Articulation) and the lowest was .02 (between number of different words and child names).

The intercorrelations (Tables 2 and 3) of the verbalization measures indicate moderately high relationships. They are, however, somewhat lower than those obtained by Templin (*3*) in her study of the language skills of children between the ages of three and eight. The moderately high positive relationships (with the exception of the moderately high negative correlation for number of one-word responses) would seem to demonstrate a common verbal factor among the verbalization measures. However, the correlations might also be interpreted in the following manner: the longer the response a child uses, the more chance he has of using more different words and the more chance he has of using sentences which are more complex in structure.

An examination of Table 3 indicates that children who utter long responses tend to use few one-word responses, as indicated by a −.52 sexes-combined correlation between length of response and number of one-word responses. A sexes-combined correlation of .68 between length of response and mean standard deviation seems to indicate that children who use long responses are more variable in the length of their uttered responses than are children who use shorter responses.

*Templin Screening Test of Articulation.* The sexes-combined correlations (Tables 1 and 2) between the Templin Test and the language and nonlanguage variables are low and often not significantly different from zero. Templin (*3*) reported correlations of .52 and .61 for five-year-old children between articulation test scores and total words used (comparable to length of response) and number of different words, respectively. Correlations of this magnitude were not found in the present study.

*Ammons Full-Range Picture Vocabulary Test.* The sexes-combined correlations between scores on the Ammons

Test and the Verbal Scale IQ and Full Scale IQ are .61 and .64, respectively. Ammons and Holmes (*1*) report a correlation of .83 between scores on their test and the Stanford-Binet MA, Form L, for children from two through five years. The higher correlation reported by Ammons and Holmes was probably a function of two things: (a) employment of an IQ test different from that used in this study and (b) computation of a correlation over a range of ages between abilities that increase with age.

*Word-Fluency Measures.* The intercorrelations between the four word-fluency measures and the correlations between the word-fluency measures and the other language and nonlanguage measures are, for the most part, positive and low. Gewirtz (*2*) reports correlations of a higher magnitude among the word-fluency measures and between the word-fluency measures and two verbalization measures (number of words and number of different words). None of his correlations, however, exceeds .50.

The interrelationships found among the various language variables were not of sufficient magnitude to warrant for predictive purposes the substitution of any one measure for any other. The relative independence of these variables suggests that until investigation indicates which measures are generally evaluated as indicative of language retardation, they should all (with the possible exception of the variability measures of verbalization and the word-fluency measures) be included as diagnostic tools of language development.

## *Summary*

Interrelationships among several language and nonlanguage variables were examined for 150 randomly selected, normal, five-year-old children (75 boys, 75 girls). Pearson product-moment correlations were computed for boys and girls separately and for the sexes combined. Comparisons of the correlations for boys and girls indicated that few of the differences were statistically significant. Very high correlations were found between the WISC Performance Scale and Verbal Scale IQ's and the Full Scale IQ. Moderately high correlations were found between the WISC Performance Scale IQ and the WISC Verbal Scale IQ; between the Ammons Test scores and the WISC Full Scale IQ; and among the various verbalization measures. Low correlations were found between socioeconomic status and the other variables; between the language measures and the Full Scale IQ; between the verbalization measures and the Ammons Test scores; between the Templin Screening Test of Articulation and the other variables; among the various word-fluency measures; and between the word-fluency measures and all of the other measures.

## *References*

1. Ammons, R. B., and Holmes, J. C., The full-range picture vocabularly test: III. Results for preschool age population. *Child Develpm.*, 20, 1949, 5-14.
2. Gewirtz, J. L., Studies in word-fluency: I. Its relation to vocabulary and mental age in young children. *J. genet. Psychol.*, 72, 1948, 165-176.
3. Templin, Mildred C., Certain language skills in children, *Inst. Child Welf., Monogr. Ser.*, No. 26. Minneapolis: Univ. Minn. Press, 1957.
4. Warner, W. L., Meeker, Marchia, and Eells, K., *Social Class in America.* Chicago: Science Res. Assoc., 1949.
5. Wechsler, D., *Manual for Wechsler Intelligence Scale for Children.* New York: Psychological Corp., 1949.
6. Wintz, H., Language skills of male and female kindergarten children. *J. Speech Hearing Res.*, 2, 1959, 377-386.

# Index to Volume 2, 1959,

# Journal of Speech and Hearing Research

Each article is indexed in three ways: by the last name of the author or authors (Author Index), by the first major word of the title (Title Index), and by subject matter (Subject Index). The following subject headings are used:

EXPERIMENTAL PHONETICS

- Acoustic Phonetics
- Auditory Feedback
- Communication Theory
- Intelligibility
- Physiological Phonetics
- Voice

GENERAL

HEARING (AUDIOLOGY)

- Audiometry (adults)
- Audiometry (children)
- Auditory Skills
- Auditory Training
- Community, Industrial, and Military Audiology
- Hearing Aids
- Physiological Acoustics
- Presbycusis
- Psychoacoustics
- Special Auditory Tests
- Speech Reading

JOURNAL DEPARTMENTS

- Book Reviews
- Graduate Theses in Speech and Hearing Research
- Letters to the Editor

LANGUAGE

- Aphasia
- Delayed Speech
- Phonetics
- Speech and Language Development

SPEECH PATHOLOGY

- Aphasia
- Articulation
- Cerebral Palsy
- Cleft Palate
- Dysarthria
- Examination and Evaluation
- Laryngectomy
- Stuttering
- Voice

AUTHOR INDEX ..... 393
TITLE INDEX ..... 395
SUBJECT INDEX ..... 397

## AUTHOR INDEX

Bedell, George N., *see* Johnson, Wendell, and others.
Bilger, Robert C., *see* Rosenblüt, Benjamin, and others.
Black, John W., Equally Contributing Frequency Bands in Intelligibility Testing: 2, March, 81-83.
Blakeley, Robert W., Erythroblastosis and Perceptive Hearing Loss: Responses of Athetoids to Test of Cochlear Function: 2, March, 5-15.
Bowser, Dolores C., *see* Sommers, Ronald K., and others.
Burk, Kenneth W., *see* Diehl, Charles F., and others.
Butler, Robert A., and Thomas Galloway, Performances of Normal-Hearing and Hard-of-Hearing Persons on the Delayed Feedback Task: 2, March, 84-90.
Byers, Vincent W., and Lewis Lieberman, Lipreading Performance and the Rate of the Speaker: 2, September, 271-276.

Carhart, Raymond, *see* Jerger, James F., and others.
Chaiklin, Joseph B., Conditioned GSR Auditory Speech Threshold: 2, September, 229-236.
Chaiklin, Joseph B., Relation Among Three Selected Auditory Speech Thresholds: 2, September, 237-243.
Chase, Richard Allen, Samuel Sutton, and Daphne First, Bibliography: Delayed Auditory Feedback: 2, June, 193-200.
Copetas, Florence G., *see* Sommers, Ronald K., and others.
Correll, Robert E., *see* Knott, John R., and others.
Curtis, James F., and James C. Hardy, Phonetic Study of Misarticulation of /r/: 2, September, 244-257.

Darley, Frederic L., Dorothy Sherman, and Gerald M. Siegel, Scaling of Abstraction Level of Single Words: 2, June, 161-167.
Davidson, G. Don, Sidetone Delay and Reading Rate, Articulation, and Pitch: 2, September, 266-270.
Diehl, Charles F., Richard White, and Kenneth W. Burk, Voice Quality and Anxiety: 2, September, 282-285.
Doerfler, Leo G., and Joan C. Kramer, Unconditioned Stimulus Strength and the Galvanic Skin Response: 2, June, 184-192.
Draper, M. H., Peter Ladefoged, and D. Whitteridge, Respiratory Muscles in Speech: 2, March, 16-27.

Feldt, Leonard S., Latin Square Design in Speech and Hearing Research: 2, September, 216-228.
First, Daphne, *see* Chase, Richard Allen, and others.
Flanagan, James L., Estimates of Intraglottal Pressure During Phonation: 2, June, 168-172.

Galloway, Thomas, *see* Butler, Robert A.
Goda, Sidney, Language Skills of Profoundly Deaf Adolescent Children: 2, December, 369-376.
Goldstein, Robert, *see* Rosenblüt, Benjamin, and others.

Hardy, James C., *see* Curtis, James F.
Harford, Earl R., and James F. Jerger, Effect of Loudness Recruitment on Delayed Speech Feedback: 2, December, 361-368.
Hess, Donald A., Pitch, Intensity, and Cleft Palate Voice Quality: 2, June, 113-125.
House, Arthur S., Note on Optimal Vocal Frequency: 2, March, 55-60.

Jerger, James F., Raymond Carhart, Tom W. Tillman, and John L. Peterson, Some Relations Between Normal Hearing for Pure Tones and for Speech: 2, June, 126-140.
Jerger, James F., *see* Harford, Earl R., . . . Loudness Recruitment. . . .
Jerger, James F., *see* Tillman, Tom W., . . . Spondee Threshold. . . .
Johnson, Wendell, Martin A. Young, Adolph L. Sahs, and George N. Bedell, Effects of Hyperventilation and Tetany on the Speech Fluency of Stutterers and Nonstutterers: 2, September, 203-215.

Knott, John R., Robert E. Correll, and Jean N. Shepherd, Frequency Analysis of Electroencephalograms of Stutterers and Nonstutterers: 2, March, 74-80.
Knower, Franklin H., Graduate Theses in Speech and Hearing Research, 1957: 2, March, 91-98.
Kramer, Joan C., *see* Doerfler, Leo G.

Ladefoged, Peter, *see* Draper, M. H., and others.
Lieberman, Lewis, *see* Byers, Vincent W.

McClintock, Colette J., *see* Sommers, Ronald K., and others.
Mysak, Edward D., Pitch and Duration Characteristics of Older Males: 2, March, 46-54.

Newby, Hayes A., *see* Ventry, Ira M.

O'Neill, John J., and Mary C. Stephens, Relationships Among Three Filmed Lipreading Tests: 2, March, 61-65.

Paul, Clara D., *see* Sommers, Ronald K., and others.
Peterson, Gordon E., Review, *Handbook of Noise Control*, Cyril M. Harris, Ed.: 2, March, 99-100.
Peterson, Gordon E., Vowel Formant Measurements: 2, June, 173-183.
Peterson, John L., *see* Jerger, James F., and others.
Powers, Gene R., *see* Spriestersbach, Duane C., Articulation Skills . . . . . Children with Cleft Palates.
Powers, Gene R., *see* Spriestersbach, Duane C., Nasality . . . . . Cleft Palate Speakers.

Quarrington, Bruce, Measures of Stuttering Adaptation: 2, June, 105-112

Rehman, Irving, *see* Snidecor, John C., and others.
Rosenblüt, Benjamin, Robert C. Bilger, and Robert Goldstein, Electrophysiologic Responses to Sound as a Function of Intensity, EEG Pattern and Sex: 2, March, 28-39.

Sahs, Adolph L., *see* Johnson, Wendell, and others.
Shelton, Ralph L., Jr., Letter to the Editor, Liaison Between ASHA and Army?: 2, March, 100-101.
Shepherd, Jean N., *see* Knott, John R., and others.
Sherman, Dorothy, *see* Darley, Frederic L., and others.
Shilling, Samuel P., *see* Sommers, Ronald K., and others.
Siegel, Gerald M., Dysphasic Speech Responses to Visual Word Stimuli: 2, June, 152-160.
Siegel, Gerald M., *see* Darley, Frederic L., and others.
Simmons, Audrey Ann, Factors Related to Lipreading: 2, December, 340-352.
Snidecor, John C., Irving Rehman, and David D. Washburn, Speech Pickup by Contact Microphone at Head and Neck Positions: 2, September, 277-281.
Sommers, Ronald K., Samuel P. Shilling, Clara D. Paul, Florence G. Copetas, Dolores C. Bowser, and Colette J. McClintock, Training Parents of Children with Functional Misarticulation: 2, September, 258-265.
Spriestersbach, Duane C., and Gene R. Powers, Articulation Skills, Velopharyngeal Closure, and Oral Breath Pressure of Children with Cleft Palates: 2, December, 318-325.
Spriestersbach, Duane C., and Gene R. Powers, Nasality in Isolated Vowels and Connected Speech of Cleft Palate Speakers: 2, March, 40-45.
Stephens, Mary C., *see* O'Neill, John J.
Stromsta, Courtney, Experimental Blockage of Phonation by Distorted Sidetone: 2, September, 286-301.
Subtelny, J. Daniel, *see* Subtelny, Joanne D.
Subtelny, Joanne D., and J. Daniel Subtelny, Intelligibility and Associated Physiological Factors of Cleft Palate Speakers: 2, December, 353-360.
Sutton, Samuel, *see* Chase, Richard Allen, and others.

Thalmann, Ruediger, Review, *Ein Binauraler Hörsynthese-Test zum Nachweis Zerebraler Hörstorungen* (A Test of Binaural Auditory Synthesis for the Detection of Cerebral Hearing Disorders), Joseph Matzker: 2, March, 99.
Tiffany, William R., Nonrandom Sources of Variation in Vowel Quality: 2, December, 305-317.
Tillman, Tom W., *see* Jerger, James F., and others.
Tillman, Tom W., and James F. Jerger, Some Factors Affecting the Spondee Threshold in Normal-Hearing Subjects: 2, June, 141-146.
Trittipoe, William J., Post-exposure Residual Effects of Low-Level Noise: 2, December, 336-339.

Ventry, Ira M., and Hayes A. Newby, Validity of the One-Frequency Screening Principle for Public School Children: 2, June, 147-151.

Wang, William S-Y., Transition and Release as Perceptual Cues for Final Plosives: 2, March, 66-73.
Washburn, David D., *see* Snidecor, John C., and others.
White, Richard, *see* Diehl, Charles F., and others.
Whitteridge, D., *see* Draper, M. H., and others.
Wingate, M. E., Calling Attention to Stuttering: 2, December, 326-335.
Winitz, Harris, Language Skills of Male and Female Kindergarten Children: 2, December, 377-386.
Winitz, Harris, Relationships Between Language and Nonlanguage Measures of Kindergarten Children: 2, December, 387-391.

Young, Martin A., *see* Johnson, Wendell, and others.

## Title Index

Articulation Skills, Velopharyngeal Closure, and Oral Breath Pressure of Children with Cleft Palates, Duane C. Spriestersbach and Gene R. Powers: 2, December, 318-325.

Bibliography: Delayed Auditory Feedback, Richard Allen Chase and others: 2, June, 193-200.
Book Reviews: *Ein Binauraler Hörsynthese-Test zum Nachweis Zerebraler Hörstorungen*, (A Test of Binaural Auditory Synthesis for the Detection of Cerebral Hearing Disorders), Joseph Matzker: 2, March, 99; *Handbook of Noise Control*, Cyril M. Harris, Ed.: 2, March, 99-100.

Calling Attention to Stuttering, M. E. Wingate: 2, December, 326-335.
Conditioned GSR Auditory Speech Threshold, Joseph B. Chaiklin: 2, September, 229-236.

Dysphasic Speech Responses to Visual Word Stimuli, Gerald M. Siegel: 2, June, 152-160.

Effect of Loudness Recruitment on Delayed Speech Feedback, Earl R. Harford and James F. Jerger: 2, December, 361-368.
Effects of Hyperventilation and Tetany on the Speech Fluency of Stutterers and Nonstutterers, Wendell Johnson and others: 2, September, 203-215.
*Ein Binauraler Hörsynthese-Test zum Nachweis Zerebraler Hörstorungen.* (A Test of Binaural Auditory Synthesis for the Detection of Cerebral Hearing Disorders), reviewed by Ruediger Thalmann: 2, March, 99 (see Book Reviews).
Electrophysiologic Responses to Sound as a Function of Intensity, EEG Pattern and Sex, Benjamin Rosenblüt and others: 2, March, 28-39.
Equally Contributing Frequency Bands in Intelligibility Testing, John W. Black: 2, March, 81-83.
Erythroblastosis and Perceptive Hearing Loss: Responses of Athetoids to Tests of Cochlear Function, Robert W. Blakeley: 2, March, 5-15.
Estimates of Intraglottal Pressure During Phonation, James L. Flanagan: 2, June, 168-172.
Experimental Blockage of Phonation by Distorted Sidetone, Courtney Stromsta: 2, September, 286-301.

Factors Related to Lipreading, Audrey Ann Simmons: 2, December, 340-352.
Formal Call for Papers, The American Speech and Hearing Association Thirty-Fifth Annual Convention, 1959: 2, March, 3-4.
Frequency Analysis of Electroencephalograms of Stutterers and Nonstutterers, John R. Knott and others: 2, March, 74-80.

Graduate Theses in Speech and Hearing Research, 1957, Franklin H. Knower: 2, March, 91-98.

*Handbook of Noise Control*, reviewed by Gordon E. Peterson: 2, March, 99-100 (see Book Reviews).

Index to Volume 2, 1959: 2, December, 392-399.
Intelligibility and Associated Physiological Factors of Cleft Palate Speakers, Joanne D. Subtelny and J. Daniel Subtelny: 2, December, 353-360.

Language Skills of Male and Female Kindergarten Children, Harris Winitz: 2, December, 377-386.
Language Skills of Profoundly Deaf Adolescent Children, Sidney Goda: 2, December, 369-376.
Latin Square Design in Speech and Hearing Research, Leonard S. Feldt: 2, September, 216-228.
Letter to the Editor: Liaison Between ASHA and Army?, Ralph L. Shelton, Jr.: 2, March, 100-101.
Liaison Between ASHA and Army?, Ralph L. Shelton, Jr.: 2, March, 100-101 (see Letter to the Editor).
Lipreading Performance and the Rate of the Speaker, Vincent W. Byers and Lewis Lieberman: 2, September, 271-276.

Measures of Stuttering Adaptation, Bruce Quarrington: 2, June, 105-112.

Nasality in Isolated Vowels and Connected Speech of Cleft Palate Speakers, Duane C. Spriestersbach and Gene R. Powers: 2, March, 40-45.
Nonrandom Sources of Variation in Vowel Quality, William R. Tiffany: 2, December, 305-317.
Note on Optimal Vocal Frequency, Arthur S. House: 2, March, 55-60.

Performances of Normal-Hearing and Hard-of-Hearing Persons on the Delayed Feedback Task, Robert A. Butler and Thomas Galloway: 2, March, 84-90.
Phonetic Study of Misarticulation of /r/, James F. Curtis and James C. Hardy: 2, September, 244-257.
Pitch and Duration Characteristics of Older Males, Edward D. Mysak: 2, March, 46-54.
Pitch, Intensity, and Cleft Palate Voice Quality, Donald A. Hess: 2, June, 113-125.
Post-exposure Residual Effects of Low-Level Noise, William J. Trittipoe: 2, December, 336-339.

Relation Among Three Selected Auditory Speech Thresholds, Joseph B. Chaiklin: 2, September, 237-243.
Relationships Among Three Filmed Lip-Reading Tests, John J. O'Neill and Mary C. Stephens: 2, March, 61-65.
Relationships Between Language and Nonlanguage Measures of Kindergarten Children, Harris Winitz: 2, December, 387-391.
Respiratory Muscles in Speech, M. H. Draper and others: 2, March, 16-27.

Scaling of Abstraction Level of Single Words, Frederic L. Darley and others: 2, June, 161-167.
Sidetone Delay and Reading Rate, Articulation, and Pitch, G. Don Davidson: 2, September, 266-270.
Some Factors Affecting the Spondee Threshold in Normal-Hearing Subjects, Tom W. Tillman and James F. Jerger: 2, June, 141-146.
Some Relations Between Normal Hearing for Pure Tones and for Speech, James F. Jerger and others: 2, June, 126-140.
Speech Pickup by Contact Microphone at Head and Neck Positions, John C. Snidecor and others: 2, September, 277-281.

Training Parents of Children with Functional Misarticulation, Ronald K. Sommers and others: 2, September, 258-265.
Transition and Release as Perceptual Cues for Final Plosives, William S-Y. Wang: 2, March, 66-73.

Unconditioned Stimulus Strength and the Galvanic Skin Response, Leo G. Doerfler and Joan C. Kramer: 2, June, 184-192.

Validity of the One-Frequency Screening Principle for Public School Children, Ira M. Ventry and Hayes A. Newby: 2, June, 147-151.
Voice Quality and Anxiety, Charles F. Diehl and others: 2, September, 282-285.
Vowel Formant Measurements, Gordon E. Peterson: 2, June, 173-183.

SUBJECT INDEX

## Experimental Phonetics

ACOUSTIC PHONETICS

Nonrandom Sources of Variation in Vowel Quality, William R. Tiffany: 2, December, 305-317.

Phonetic Study of Misarticulation of /r/, James F. Curtis and James C. Hardy: 2, September, 244-257.

Transition and Release as Perceptual Cues for Final Plosives, William S-Y. Wang: 2, March, 66-73.

Vowel Formant Measurements, Gordon E. Peterson: 2, June, 173-183.

AUDITORY FEEDBACK

Bibliography: Delayed Auditory Feedback, Richard Allen Chase and others: 2, June, 193-200.

Effect of Loudness Recruitment on Delayed Speech Feedback, Earl R. Harford and James F. Jerger: 2, December, 361-368.

Experimental Blockage of Phonation by Distorted Sidetone, Courtney Stromsta: 2, September, 286-301.

Performances of Normal-Hearing and Hard-of-Hearing Persons on the Delayed Feedback Task, Robert A. Butler and Thomas Galloway: 2, March, 84-90.

Sidetone Delay and Reading Rate, Articulation, and Pitch, G. Don Davidson: 2, September, 266-270.

COMMUNICATION THEORY

INTELLIGIBILITY

Equally Contributing Frequency Bands in Intelligibility Testing, John W. Black: 2, March, 81-83.

PHYSIOLOGICAL PHONETICS

Estimates of Intraglottal Pressure During Phonation, James L. Flanagan: 2, June, 168-172.

Respiratory Muscles in Speech, M. H. Draper and others: 2, March, 16-27.

Speech Pickup by Contact Microphone at Head and Neck Positions, John C. Snidecor and others: 2, September, 277-281.

VOICE

Note on Optimal Vocal Frequency, Arthur S. House: 2, March, 55-60.

Pitch and Duration Characteristics of Older Males, Edward D. Mysak: 2, March, 46-54.

Speech Pickup by Contact Microphone at Head and Neck Positions, John C. Snidecor and others: 2, September, 277-281.

## General

Index to Volume 2, 1959: 2, December, 392-399.

Latin Square Design in Speech and Hearing Research, Leonard S. Feldt: 2, September, 216-228.

## Hearing (Audiology)

AUDIOMETRY (ADULTS)

Conditioned GSR Auditory Speech Threshold, Joseph B. Chaiklin: 2, September, 229-236.

Electrophysiologic Responses to Sound as a Function of Intensity, EEG Pattern and Sex, Benjamin Rosenblüt and others: 2, March, 28-39.

Equally Contributing Frequency Bands in Intelligibility Testing, John W. Black: 2, March, 81-83.

Relation Among Three Selected Auditory Speech Thresholds, Joseph B. Chaiklin: 2, September, 237-243.

Some Factors Affecting the Spondee Threshold in Normal-Hearing Subjects, Tom W. Tillman and James F. Jerger: 2, June, 141-146.
Some Relations Between Normal Hearing for Pure Tones and for Speech, James F. Jerger and others: 2, June, 126-140.

AUDIOMETRY (CHILDREN)

Conditioned GSR Auditory Speech Threshold, Joseph B. Chaiklin: 2, September, 229-236.
Electrophysiologic Responses to Sound as a Function of Intensity, EEG Pattern and Sex, Benjamin Rosenblüt and others: 2, March, 28-39.
Relation Among Three Selected Auditory Speech Thresholds, Joseph B. Chaiklin: 2, September, 237-243.
Validity of the One-Frequency Screening Principle for Public School Children, Ira M. Ventry and Hayes A. Newby: 2, June, 147-151.

AUDITORY SKILLS

AUDITORY TRAINING

COMMUNITY, INDUSTRIAL, AND MILITARY AUDIOLOGY

HEARING AIDS

PHYSIOLOGICAL ACOUSTICS

Electrophysiologic Responses to Sound as a Function of Intensity, EEG Pattern and Sex, Benjamin Rosenblüt and others: 2, March, 28-39.
Post-exposure Residual Effects of Low-Level Noise, William J. Trittipoe: 2, December, 336-339.

PRESBYCUSIS

PSYCHOACOUSTICS

Conditioned GSR Auditory Speech Threshold, Joseph B. Chaiklin: 2, September, 229-236.
Equally Contributing Frequency Bands in Intelligibility Testing, John W. Black: 2, March, 81-83.
Post-exposure Residual Effects of Low-Level Noise, William J. Trittipoe: 2, December, 336-339.
Relation Among Three Selected Auditory Speech Thresholds, Joseph B. Chaiklin: 2, September, 237-243.
Some Factors Affecting the Spondee Threshold in Normal-Hearing Subjects, Tom W. Tillman and James F. Jerger: 2, June, 141-146.
Some Relations Between Normal Hearing for Pure Tones and for Speech, James F. Jerger and others: 2, June, 126-140.

SPECIAL AUDITORY TESTS

Effect of Loudness Recruitment on Delayed Speech Feedback, Earl R. Harford and James F. Jerger: 2, December, 361-368.
Electrophysiologic Responses to Sound as a Function of Intensity, EEG Pattern and Sex, Benjamin Rosenblüt and others: 2, March, 28-39.
Performances of Normal-Hearing and Hard-of-Hearing Persons on the Delayed Feedback Task, Robert A. Butler and Thomas Galloway: 2, March, 84-90.
Unconditioned Stimulus Strength and the Galvanic Skin Response, Leo G. Doerfler and Joan C. Kramer: 2, June, 184-192.

SPEECH READING

Factors Related to Lipreading, Audrey Ann Simmons: 2, December, 340-352.
Lipreading Performance and the Rate of the Speaker, Vincent W. Byers and Lewis Lieberman: 2, September, 271-276.
Relationships Among Three Filmed Lip-Reading Tests, John J. O'Neill and Mary C. Stephens: 2, March, 61-65.

## Journal Departments

ANNOUNCEMENTS

Formal Call for Papers, The American Speech and Hearing Association Thirty-Fifth Annual Convention, 1959: 2, March, 3-4.

BOOK REVIEWS

*Ein Binauraler Hörsynthese-Test zum Nachweis Zerebraler Hörstorungen* (A Test of Binaural Auditory Synthesis for the Detection of Cerebral Hearing Disorders), Joseph Matzker, reviewed by Ruediger Thalmann: 2, March, 99.

*Handbook of Noise Control*, Cyril M. Harris, Ed., reviewed by Gordon E. Peterson: 2, March, 99-100.

GRADUATE THESES IN SPEECH AND HEARING RESEARCH

Graduate Theses in Speech and Hearing Research, 1957, Franklin H. Knower: 2, March, 91-98.

LETTERS TO THE EDITOR

Liaison Between ASHA and Army?, Ralph L. Shelton, Jr.: 2, March, 100-101.

## Language

APHASIA

Dysphasic Speech Responses to Visual Word Stimuli, Gerald M. Siegel: 2, June, 152-160.

Scaling of Abstraction Level of Single Words, Frederic L. Darley and others: 2, June, 161-167.

DELAYED SPEECH

PHONETICS

SPEECH AND LANGUAGE DEVELOPMENT

Language Skills of Male and Female Kindergarten Children, Harris Winitz: 2, December, 377-386.

Language Skills of Profoundly Deaf Adolescent Children, Sidney Goda: 2, December, 369-376.

Relationships Between Language and Nonlanguage Measures of Kindergarten Children, Harris Winitz: 2, December, 387-391.

## Speech Pathology

APHASIA

Dysphasic Speech Responses to Visual Word Stimuli, Gerald M. Siegel: 2, June, 152-160.

ARTICULATION

Phonetic Study of Misarticulation of /r/, James F. Curtis and James C. Hardy: 2, September, 244-257.

Training Parents of Children with Functional Misarticulation, Ronald K. Sommers and others: 2, September, 258-265.

CEREBRAL PALSY

Erythroblastosis and Perceptive Hearing Loss: Responses of Athetoids to Tests of Cochlear Function, Robert W. Blakeley: 2, March, 5-15.

CLEFT PALATE

Articulation Skills, Velopharyngeal Closure, and Oral Breath Pressure of Children with Cleft Palates, Duane C. Spriestersbach and Gene R. Powers: 2, December, 318-325.

Intelligibility and Associated Physiological Factors of Cleft Palate Speakers, Joanne D. Subtelny and J. Daniel Subtelny: 2, December, 353-360.

Nasality in Isolated Vowels and Connected Speech of Cleft Palate Speakers, Duane C. Spriestersbach and Gene R. Powers: 2, March, 40-45.

Pitch, Intensity, and Cleft Palate Voice Quality, Donald A. Hess: 2, June, 113-125.

DYSARTHRIA

EXAMINATION AND EVALUATION

LARYNGECTOMY

STUTTERING

Calling Attention to Stuttering, M. E. Wingate: 2, December, 326-335.

Effects of Hyperventilation and Tetany on the Speech Fluency of Stutterers and Nonstutterers, Wendell Johnson and others: 2, September, 203-215.

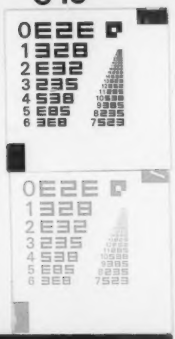
Frequency Analysis of Electroencephalograms of Stutterers and Nonstutterers, John R. Knott and others: 2, March, 74-80.

Measures of Stuttering Adaptation, Bruce Quarrington: 2, June, 105-112.

VOICE

Pitch, Intensity, and Cleft Palate Voice Quality, Donald A. Hess: 2, June, 113-125.

Voice Quality and Anxiety, Charles F. Diehl and others: 2, September, 282-285.

1592

C15